STUDIES IN HEALTH TECHNOLOGY AND INFORMATICS 263

Applied Interdisciplinary Theory in Health Informatics

A Knowledge Base for Practitioners

Editors: Philip Scott
Nicolette de Keizer
Andrew Georgiou

IOS Press

APPLIED INTERDISCIPLINARY THEORY IN HEALTH INFORMATICS

Studies in Health Technology and Informatics

International health informatics is driven by developments in biomedical technologies and medical informatics research that are advancing in parallel and form one integrated world of information and communication media and result in massive amounts of health data. These components include genomics and precision medicine, machine learning, translational informatics, intelligent systems for clinicians and patients, mobile health applications, data-driven telecommunication and rehabilitative technology, sensors, intelligent home technology, EHR and patient-controlled data, and Internet of Things.

Studies in Health Technology and Informatics (HTI) series was started in 1990 in collaboration with EU programmes that preceded the Horizon 2020 to promote biomedical and health informatics research. It has developed into a highly visible global platform for the dissemination of original research in this field, containing more than 250 volumes of high-quality works from all over the world.

The international Editorial Board selects publications with relevance and quality for the field. All contributions to the volumes in the series are peer reviewed.

The HTI series is indexed by MEDLINE/PubMed; Web of Science: Conference Proceedings Citation Index – Science (CPCI-S) and Book Citation Index – Science (BKCI-S); Google Scholar; Scopus; EMCare.

Volume 263

Recently published in this series

Vol. 262 J. Mantas, A. Hasman, P. Gallos, A. Kolokathi, M.S. Househ and J. Liaskos (Eds.), Health Informatics Vision: From Data via Information to Knowledge

Vol. 261 B. Blobel and M. Giacomini (Eds.), pHealth 2019 – Proceedings of the 16th International Conference on Wearable Micro and Nano Technologies for Personalized Health – 10–12 June 2019, Genoa, Italy

Vol. 260 D. Hayn, A. Eggerth and G. Schreier (Eds.), dHealth 2019 – From eHealth to dHealth – Proceedings of the 13th Health Informatics Meets Digital Health Conference

Vol. 259 T. Bürkle, M. Lehmann, K. Denecke, M. Sariyar, S. Bignens, E. Zetz and J. Holm (Eds.), Healthcare of the Future – Bridging the Information Gap – 5 April 2019, Biel/Bienne, Switzerland

ISSN 0926-9630 (print)
ISSN 1879-8365 (online)

Applied Interdisciplinary Theory in Health Informatics

A Knowledge Base for Practitioners

Edited by

Philip Scott

Reader in Health Informatics, Centre for Healthcare Modelling & Informatics, University of Portsmouth, UK

Nicolette de Keizer

Professor of Medical Informatics, Department of Medical Informatics, Amsterdam UMC, University of Amsterdam, Amsterdam Public Health Institute, The Netherlands

and

Andrew Georgiou

Professor of Health Informatics, Centre for Health Systems and Safety Research, Macquarie University, Sydney, New South Wales, Australia

IOS Press

Amsterdam • Berlin • Washington, DC

ISBN 978-1-61499-990-4 (print)
ISBN 978-1-61499-991-1 (online)
Library of Congress Control Number: 2019946174
doi: 10.3233/SHTI263

Publisher
IOS Press BV
Nieuwe Hemweg 6B
1013 BG Amsterdam
Netherlands
fax: +31 20 687 0019
e-mail: order@iospress.nl

For book sales in the USA and Canada:
IOS Press, Inc.
6751 Tepper Drive
Clifton, VA 20124
USA
Tel.: +1 703 830 6300
Fax: +1 703 830 2300
sales@iospress.com

PRINTED IN THE NETHERLANDS

Foreword

Charles P. FRIEDMAN[a,1]

[a]*Josiah Macy Jr. Professor of Medical Education*
Chair, Department of Learning Health Sciences, Medical School
Professor of Information and Public Health
University of Michigan
Editor-in-Chief: Learning Health Systems

This volume is an enormously important step, in a vital direction, for the field of health informatics.

Few would argue with a statement that health informatics (HI) is an applied science. Given almost universal consensus around that assertion, HI researchers can approach their work from three positions. The first is what might be called pure empiricism, holding to the premise that HI does not need theories. Knowledge in the field can cumulate through results of empirical studies, grounded in real world experiences, that reference and build upon each other. The second position is what might be called theoretical derivativism. As an applied field, HI does not require theories of its own, and it suffices to borrow theories from the widely acknowledged basic sciences such as mathematics, computer science, philosophy, psychology, and sociology. Supporters of the second position would also assert that HI could borrow theory from more mature, but also applied, fields such as policy science, information science, and decision science. The third position – the most radical and, to me, most appealing – is that HI could, over time, evolve its own theoretical bases and assume a place alongside other, more mature applied sciences.

This volume importantly and emphatically reveals the inadequacy of the first position, argues elegantly for and enables the second, and stimulates us to contemplate the third. It accomplishes this by describing in depth a strong candidate set of theories that do, could, or should underpin research in HI. These contributions support a statement that HI must be evidence-based and, perhaps more important, a statement that HI should also be prediction-based. Prediction-based HI research studies are grounded in models of the world that exist at higher levels of abstraction than the measures used in the study itself. These abstract models are then applied to describe *expected* study results, in advance of any data collection and without resort to the results of other studies. In the pages that follow, the reader will experience an invaluable catalog of perspectives that can achieve not only evidence-based informatics but also prediction-based informatics.

And peeking over the horizon, using the perspectives offered in this volume, everyone in the field, can begin to visualize what a theoretical basis unique to health informatics might be. Will this take the form of a synthesis of the fields described herein, or will some new perspectives, perhaps suggested by what is not included in this finite volume,

[1] Corresponding Author: Prof Charles Friedman, E-mail: cpfried@umich.edu

come into view? There are those who believe – myself among them – that there is something scientifically unique about informatics. It is perhaps found in the way informatics activities must span the life cycle of information resources: from design to test to deployment to evaluation, and re-design based on evaluation. It is perhaps found in the admixture of the promise of computation and the well-being of humanity that attracted so many creative people to this remarkable field in the first place. Whatever this unique theoretical perspective might turn out to be, if anything, the journey is undoubtedly worth taking, and in many ways, it begins with this book.

Preface

Philip J. SCOTT[a,1], Nicolette F. DE KEIZER[b] and Andrew GEORGIOU[c]
[a]*Centre for Healthcare Modelling & Informatics, University of Portsmouth, UK*
[b]*Department of Medical Informatics, Amsterdam UMC, University of Amsterdam, Amsterdam Public Health Institute, The Netherlands*
[c]*Centre for Health Systems and Safety Research, Macquarie University, Sydney, New South Wales, Australia*

1. Purpose

Kurt Lewin, the pioneer of social psychology, famously said that 'there is nothing more practical than a good theory' [1]. We agree and hope that readers of this book will come to share this view. Our aim is to provide a scientific knowledge base to support education, research and implementation. The editors came together as proponents of evidence-based health informatics within the European Federation of Medical Informatics (EFMI) Working Group on Health IT Evaluation and the International Medical Informatics Association (IMIA) Working Group on Technology Assessment and Quality Development [2]. We have the shared belief that theory is insufficiently considered in our field along with the collegiate aim to improve the status quo. We want to move theory from a niche interest to a core concern of health informatics, to contribute to the maturity of the discipline and above all to improve care by effective health IT interventions. Specifically, this book was motivated by the outcome of a workshop at Medinfo 2015 that called for a "theory toolbox", as elaborated in a paper at Medical Informatics Europe 2016 [3].

There are distinct audiences and corresponding benefits from taking a theoretically-informed approach to health informatics. For implementers, application of theory can help **adoption of best practice** and work towards **demonstrating improved outcomes** of health informatics interventions. For researchers and evaluators, knowledge of theory can help to identify **gaps in knowledge** and hence **prioritise, justify and guide research and evaluation** where they are most needed. For educators, using theory can **instil a scientific approach** in their students. Importantly, we believe that all of these groups can be termed "practitioners" of health informatics (as our book title suggests).

Overall, the purpose of the book is to move forward the agenda of evidence-based health informatics [4] by emphasising theory-informed work that aims to "enrich our … understanding of this uniquely complex field" [5]. We have not set out to offer an exhaustive or comprehensive coverage of theory in health informatics. As our final chapter elaborates, we know that there are important topics that we have not been able to include in this volume. However, we do believe that this book discusses some of the most important and commonly used theories relevant to health informatics and that this collection marks a significant milestone on the journey. We want this book to constitute a first iteration of a consolidated knowledge base that can advance the science of our field.

[1] Corresponding Author: Philip Scott, E-mail: philip.scott@port.ac.uk

To introduce the textbook, let us first clarify its scope: What is "theory"? What is "health informatics"? Why "interdisciplinary" theory?

2. Varying Perceptions of "Theory"

Our recurring experience in the production of this book has been the diversity of opinion about what exactly "theory" means. From our initial discussions, through the process of defining scope, commissioning chapters, inviting peer reviews and appraising author revisions we have repeatedly had to step back to reflect and question our own common understanding and that of our numerous contributors.

We found the Nilsen theory categories [6] a helpful anchor point to specify various types of theory (discussed further in the next chapter) and we cited the Nilsen paper in our brief to authors. Even then, we found that authors and reviewers did not always apply the categories uniformly or in line with our own editorial perceptions. We think that this tells us something about our field. While there will inevitably be some continuing academic pedantry and diverse schools of thought around particular concepts in the metaphysics of epistemology and methodology, we were surprised by the degree of divergence. Of course, there is also "theory" in the more generic sense of "a body of knowledge", such as "social theory" or "economic theory", but that is a different level of abstraction to our subject matter of *specific* theories (though not always a distinction that can be neatly maintained). Our experience is that health informatics is not a field that has a recognized common language to talk about its foundational ideas. Hence, we recall Kuhn's seminal discussion of the progression of science and his reference to the "paradigm" of a discipline [7] and must question whether health informatics is yet a "mature" science. We return to this discussion in our final chapter.

There are "soft" and "hard" definitions of theory. To some extent these may reflect their respective disciplinary research tradition as primarily qualitative or quantitative in approach, but that is by no means a fixed rule and in any case is not unique to health informatics. The interdisciplinary nature of health informatics necessarily brings together people with varying cultural and practice norms, as we discuss further below. A "soft" definition might be that a theory comprises a hypothesis or a set of general principles within a defined conceptual model (a "determinant framework" in Nilsen's terminology). A "hard" definition might be that a theory will make testable and quantitatively measurable predictions (a "classic theory" in Nilsen's description). If we can accept a spectrum of theory types that incorporates both "soft" and "hard" definitions, then we have an approach that is broad enough to include everything from a theorised qualitative explanation (such as a "grounded theory") through to equations that predict the relative clinical utility of particular laboratory tests. For this textbook, we have pragmatically adopted a flexible and inclusive view of theory. We asked chapter authors to work with the theory description: "*abstract enough to permit generalization, but concrete enough to permit testing*". After Merton, we characterised these as "middle-range" [8] theories, not grand "theories of everything" but "special theories from which to derive hypotheses that can be empirically investigated". By "testability" and "empirical investigation", we mean simply that the given theory can be shown to have made a difference in some aspect of a health informatics lifecycle: design, validation, verification, implementation and evaluation.

3. Definitions of Health Informatics

There is still no single universally agreed definition of health informatics, but we now seem to have a converging set of ideas. Although the principal professional societies still use the older and narrower term "medical informatics" in their organizational names (e.g. EFMI, IMIA and the American Medical Informatics Association, AMIA), they each promote more inclusive wording in their official publications. Early definitions of *medical* informatics were:

- "the field that concerns itself with the cognitive, information processing, and communication tasks of medical practice, education, and research, including the information science and the technology to support these tasks" [9]
- "the scientific discipline concerned with the systematic processing of data, information and knowledge in medicine and health care" [10]

Whereas IMIA prefers the phrase "biomedical and health informatics" (BMHI) [11], AMIA favours the term "biomedical informatics" (BMI), which it defines [12] as:

- "the interdisciplinary field that studies and pursues the effective uses of biomedical data, information, and knowledge for scientific inquiry, problem solving and decision making, motivated by efforts to improve human health".

In this definition, BMI is the "scientific core" that is applied in the domains of *bio*-informatics and *imaging* informatics, *health* informatics (comprising *clinical* and *public health* informatics) and *translational* informatics. It has been argued that the more holistic term "biopsychosocial" would be a better adjective than "biomedical" [13], but in its World Health Organization definition the global term "health" subsumes all aspects of physical, mental and social well-being [14]. Therefore, we use the term "health informatics" as a simple and inclusive descriptor to cover both BMHI and BMI.

We find the AMIA definition particularly helpful in its articulation of the three "foundational domains" of health informatics: health science, information science, and social science and their various overlaps (see Figure 1, from [15]). We have used this model to structure the content of this textbook around the major subject areas.

4. Meaning and Importance of Interdisciplinarity in Health Informatics

Whatever label we choose to adopt for our field, it is unquestionably "interdisciplinary" as noted in the AMIA definition. Reflecting the three foundational domains, it is always the case that health informatics needs both healthcare and information science knowledge and skills. Increasingly often, the importance of the social sciences is also recognized. Interdisciplinary is defined as "contributing to or benefiting from two or more disciplines" [16] and is helpfully distinguished from "multidisciplinary" and "transdisciplinary" in the following summary [17]:

- Multidisciplinarity draws on knowledge from different disciplines but stays within their boundaries. Interdisciplinarity analyzes, synthesizes and harmonizes links between disciplines into a coordinated and coherent whole. Transdisciplinarity integrates the natural, social and health sciences in a humanities context, and transcends their traditional boundaries.

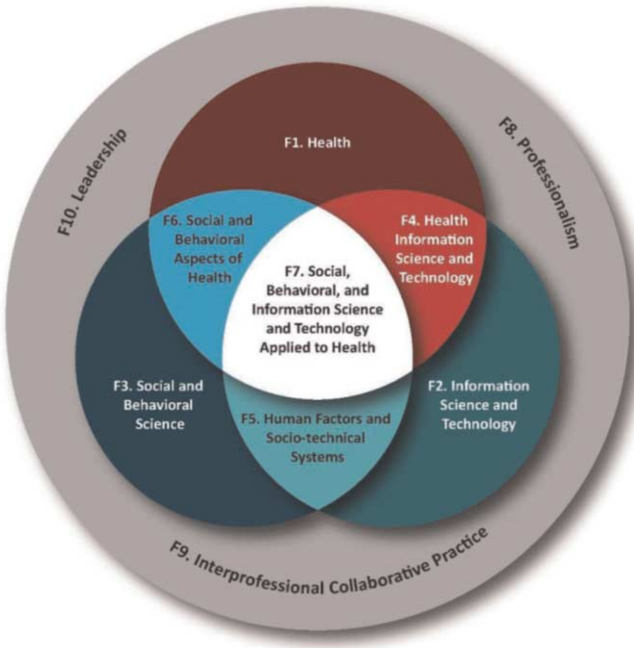


Figure 1. Foundational domains of applied health informatics.
(Reproduced from [15] by permission of Oxford University Press.)

We do not suggest that health informatics cannot be transdisciplinary or multi-disciplinary at times. However, we do propose that interdisciplinarity is the term that best describes most good health informatics work today. We do not want to stay within disciplinary boundaries, but we do aspire to offer coherent synthesis across disciplines. Transdisciplinarity may sometimes be attained, but we suggest it is perhaps too high a goal and not necessarily a priority for most resource-limited health informatics work [18,19].

5. Structure of the Book

We set our overall learning aim for the textbook as: What theories have been applied in health informatics and what difference have they made? The specific objectives were:

- To show where and how interdisciplinary theories have been applied in health informatics
- To identify theory developed specifically within health informatics
- To highlight where further work is necessary to develop theory-based approaches.

We undertook a consultative exercise with IMIA and EFMI Working Group members and potential chapter authors about relevant topics to feature, based around the three foundational domains of health science, information science and social science. When we invited authors to submit chapters, we proposed a standard structure to aid navigation and so that each chapter could be used as a standalone entity for educational use.

There is inevitably some debate about which theory belongs to which domain of knowledge, but we have ended up with sections that address only two of the three foundational domains in the AMIA model. The omission of theories from health sciences is not by design, as we discuss further in the final chapter. Section 1 deals with theories from information science and technology, such as general system theory, technology adoption models and Shannon's information theory. Section 2 addresses theories from the social and psychological sciences such as distributed cognition, resilience theory and normalisation process theory. In Section 3, we offer two kinds of synthesis. Firstly, we consider the ambitious framework described by Greenhalgh and Abimbola that aims to integrate several theoretical approaches to the adoption and sustainability of health informatics interventions. Secondly, as editors we offer our own overview of theory within the overall health informatics body of knowledge and propose a research agenda. In this chapter we highlight where further work is necessary to develop theory-based approaches and mature the health informatics discipline.

6. Suggested Use in Teaching

We suggest that the specified learning objectives in each chapter might be used to construct a teaching plan for a given lecture or seminar. Students could be assigned, individually or in small groups, to produce reflective reports based upon directed reading of one or more of the chapter references. The questions for reflection at the end of the chapter might be featured in coursework or in interactive seminars. Students could be asked to find additional illustrations of the theory's usage in health informatics, contrasting examples in other fields, or how alternative theories were applied in analogous scenarios. We encourage reflection on how the use (or non-use) of theory can explain relative success or failure in health informatics and on the maturity of theory in the field. Doctoral students may like to study the gaps or weaknesses in theory: where can new contributions be made?

Acknowledgements

The editors gratefully acknowledge all our colleagues who gave formative advice in the planning of this book and, of course, all our chapter authors and peer reviewers.

References

[1] K. Lewin, *Field theory in social science: selected theoretical papers*, Harper, New York, NY, 1951.
[2] M. Rigby, E. Ammenwerth, M.-C. Beuscart-Zephir, J. Brender, H. Hyppönen, S. Melia, P. Nykänen, J. Talmon, and N. de Keizer, Evidence Based Health Informatics: 10 Years of Efforts to Promote the Principle, *Yearb Med Inform* **34** (2013), 46.
[3] P.J. Scott, A. Georgiou, H. Hypponen, C.K. Craven, M. Rigby, and J. Brender McNair, Theoretical Foundations for Evidence-Based Health Informatics: Why? How?, *Stud Health Technol Inform* **228** (2016), 614–618.
[4] E. Ammenwerth and M. Rigby, eds., *Evidence-based Health Informatics: Promoting safety and efficiency through scientific methods and ethical policy*, IOS Press, Amsterdam, 2016.
[5] T. Greenhalgh, H.W. Potts, G. Wong, P. Bark, and D. Swinglehurst, Tensions and paradoxes in electronic patient record research: a systematic literature review using the meta-narrative method, *Milbank Q* **87** (2009), 729–788.

[6] P. Nilsen, Making sense of implementation theories, models and frameworks, *Implement Sci* **10** (2015), 53.
[7] T.S. Kuhn, *The structure of scientific revolutions*, University of Chicago, Chicago, IL, 1970.
[8] R.K. Merton, On sociological theories of the middle range [1949], in: *Social Theory and Social Structure,* R.K. Merton, ed., Simon & Schuster, New York, 1949, pp. 39–53.
[9] R.A. Greenes and E.H. Shortliffe, Medical Informatics: An Emerging Academic Discipline and Institutional Priority, *JAMA* **263** (1990), 1114–1120.
[10] J. van Bemmel and M. Musen, *Handbook of Medical Informatics*, Springer-Verlag, Heidelberg, 1997.
[11] J. Mantas, E. Ammenwerth, G. Demiris, A. Hasman, R. Haux, W. Hersh, E. Hovenga, K.C. Lun, H. Marin, F. Martin-Sanchez, and G. Wright, Recommendations of the International Medical Informatics Association (IMIA) on Education in Biomedical and Health Informatics. First Revision, *Methods Inf Med* **49** (2010), 105–120.
[12] C.A. Kulikowski, E.H. Shortliffe, L.M. Currie, P.L. Elkin, L.E. Hunter, T.R. Johnson, I.J. Kalet, L.A. Lenert, M.A. Musen, J.G. Ozbolt, J.W. Smith, P.Z. Tarczy-Hornoch, and J.J. Williamson, AMIA Board white paper: definition of biomedical informatics and specification of core competencies for graduate education in the discipline, *J Am Med Inform Assoc* **19** (2012), 931–938.
[13] S. de Lusignan, What is primary care informatics?, *Journal of the American Medical Informatics Association* **10** (2003), 304–309.
[14] World Health Organization, Preamble to the Constitution of WHO, in: *Official Records of WHO*, 1946, p. 100.
[15] A.L. Valenta, E.S. Berner, S.A. Boren, G.J. Deckard, C. Eldredge, D.B. Fridsma, C. Gadd, Y. Gong, T. Johnson, J. Jones, E.L. Manos, K.T. Phillips, N.K. Roderer, D. Rosendale, A.M. Turner, G. Tusch, J.J. Williamson, and S.B. Johnson, AMIA Board White Paper: AMIA 2017 core competencies for applied health informatics education at the master's degree level, *J Am Med Inform Assoc* **25** (2018), 1657–1668.
[16] Oxford English Dictionary, "*interdisciplinary, adj.*", Oxford University Press, 1989.
[17] B.C. Choi and A.W. Pak, Multidisciplinarity, interdisciplinarity and transdisciplinarity in health research, services, education and policy: 1. Definitions, objectives, and evidence of effectiveness, *Clin Invest Med* **29** (2006), 351–364.
[18] J. Fawcett, Thoughts about multidisciplinary, interdisciplinary, and transdisciplinary research, *Nurs Sci Q* **26** (2013), 376–379.
[19] A. O'Cathain, E. Murphy, and J. Nicholl, Multidisciplinary, interdisciplinary, or dysfunctional? Team working in mixed-methods research, *Qual Health Res* **18** (2008), 1574–1585.

Acknowledgements

The editors would like to gratefully acknowledge the expert assistance of our many colleagues in formative advice and detailed peer review of submitted chapters.

Ammenwerth, Elske
Braithwaite, Jeffrey
Callen, Joanne
Chlebikova, Janka
Cornet, Ronald
Day, Karen
Evans, Ruth
Francis, Jill
Glidewell, Liz
Göran, Henriks
Greenhalgh, Trish
Grimshaw, Jeremy
Gude, Wouter
Hannan, Terry
Heitmann, Kai
Jaspers, Monique
Koppel, Ross
Krause, Paul
Lichtner, Valentina
Mackintosh, Nicola
Moreira Marquez, Marta
Murray, Elizabeth
Nakkas, Haythem
Neher, Margit
Patey, Andrea
Perry, Mark
Peek, Niels
Pemberton, Lyn
Roberts, Karlene
Ruiter, Rob
Sarcevic, Aleksandra
Sujan, Mark-Alexander
Sullivan, Frank
Ubbink, Dirk
Whetton, Sue
Rampton, Suzanne
Randell, Rebecca
Roberts, Karlene
Sujan, Mark-Alexander
Thun, Sylvia
van der Veer, Sabine
Whetton, Sue
Williams, Richard
Wong, Zoie Shui-Yee

Contents

Foreword v
Charles P. Friedman

Preface vii
Philip J. Scott, Nicolette F. de Keizer and Andrew Georgiou

Acknowledgements xiii

Chapter 1. The Need for Theory to Inform Clinical Information Systems and Professionalise the Health Informatics Discipline 1
Jeremy C. Wyatt

Part 1. Information Science and Technology Theories

Chapter 2. General System Theory and the Use of Process Mining to Improve Care Pathways 11
Owen Johnson

Chapter 3. Information Theory and Medical Decision Making 23
Paul Krause

Chapter 4. Assessing Technology Success and Failure Using Information Value Chain Theory 35
Enrico Coiera

Chapter 5. Linking Activity Theory with User Centred Design: A Human Computer Interaction Framework for the Design and Evaluation of mHealth Interventions 49
Alice Good and Omobolanle Omisade

Chapter 6. Technology Acceptance Models in Health Informatics: TAM and UTAUT 64
Elske Ammenwerth

Part 2. Social and Psychological Theories

Chapter 7. Distributed Cognition: Understanding Complex Sociotechnical Informatics 75
Dominic Furniss, Sara Garfield, Fran Husson, Ann Blandford and Bryony Dean Franklin

Chapter 8. Using Actor-Network Theory to Study Health Information Technology Interventions 87
Kathrin Cresswell

Chapter 9. Collective Mindfulness and Processes of Sensemaking in Health IT Implementation 98
Valentina Lichtner and Johanna I. Westbrook

Chapter 10. Fostering Shared Decision Making with Health Informatics Interventions Based on the Boosting Framework 109
Marieke de Vries, Jesse Jansen, Julia van Weert and Rob Holland

Chapter 11. Junior Doctor Communication Systems and the Deterioration Communication Management Theory 122
Janet Liang, Jim Warren, Martin Orr and Karen Day

Chapter 12. Resilient Health Care: A Determinant Framework for Understanding Variation in Everyday Work and Designing Sustainable Digital Health Systems 134
Robyn Clay-Williams and Jeffrey Braithwaite

Chapter 13. Health Behaviour Theory in Health Informatics: Support for Positive Change 146
Stephanie Medlock and Jeremy C. Wyatt

Chapter 14. Control Theory to Design and Evaluate Audit and Feedback Interventions 159
Wouter T. Gude and Niels Peek

Chapter 15. Implementing and Embedding Health Informatics Systems – Understanding Organizational Behaviour Change Using Normalization Process Theory (NPT) 171
Mike Bracher and Carl R. May

Part 3. Synthesis

Chapter 16. The NASSS Framework – A Synthesis of Multiple Theories of Technology Implementation 193
Trisha Greenhalgh and Seye Abimbola

Chapter 17. Reflecting and Looking to the Future: What Is the Research Agenda for Theory in Health Informatics? 205
Philip J. Scott, Nicolette F. de Keizer and Andrew Georgiou

Subject Index 219

Author Index 221

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

doi:10.3233/SHTI190105

The Need for Theory to Inform Clinical Information Systems and Professionalise the Health Informatics Discipline

Jeremy C. WYATT[a,1]
[a] *Professor of Digital Healthcare and Director, Wessex Institute of Health Research, Faculty of Medicine, University of Southampton, UK*

Abstract. This chapter introduces the idea of theories in health informatics, defines what we mean by theory and distinguishes theories from models, frameworks and predictive principles. After explaining why theories and predictive principles are needed to help us professionalize our discipline, the chapter offers five criteria for a successful predictive principle, discusses how to evaluate predictive principles and theories and links this with the emerging field of evidence-based health informatics. The chapter concludes with three actions needed to move the discipline of theory-based health informatics forward.

Keywords. Theory, Professionalism, Evaluation, Scientific methods, Health informatics

Learning objectives

After reading this chapter, the reader will be able to:

1. Define “theory”, know where to locate relevant theories and understand what types of theory are relevant to health informatics.
2. Describe the importance of predictive principles and theories in advancing the health informatics discipline, health informatics research and in educating students and practitioners.
3. Explain the importance of theory in developing usable, effective health information systems, and the relevance of theory to procurement decisions.

1. Introduction: What are theories, and why do we need them in health informatics ?

Nilsen [1] has written a valuable contribution about the nature of theory in the related field of implementation science which I believe should also inform our work in health informatics (HI). She conducted a careful review of the types and uses of theory and of the related concepts, models and frameworks [1] in her discipline, with two main findings.

[1] Corresponding Author: Jeremy Wyatt, E-mail: j.c.wyatt@soton.ac.uk

First, she distinguished carefully between a ***theory*** (a set of analytical principles designed to structure our observations, understanding and explanation of a phenomenon in the world; a theory must be explanatory and predictive as well as descriptive), a ***framework*** (a structure, overview or plan consisting of various elements, concepts, or variables and their relationships, intended to describe a phenomenon) and a ***model*** (a deliberate simplification of a phenomenon or part of a phenomenon, typically descriptive). Unlike theories, frameworks and models do not specify the mechanisms of change. Second, she identified five important categories of theory, model or framework, based largely on their origin and intended use [1]:

1. Process models: these typically specify the steps to be followed to achieve some goal. A historic example relevant to health informatics could be the waterfall model for software engineering (Figure 1)
2. Determinant frameworks: a list of key determinants (e.g. barriers and enablers of change) and their relationships that may influence project outcome. An example relevant to health informatics is Schneiderman's checklist of eight user interface features associated with high usability (see Box 1) [2]
3. Classic theories: these are predictive theories that arise from external disciplines (e.g. psychology, sociology or management science) that can assist understanding and / or explanation. An example relevant to health informatics is Michie's COM-B theory: Behaviour change requires Capability, Opportunity and Motivation [3]
4. Implementation theories: theories that arise from within the implementation science discipline that can assist understanding and / or explanation. Note that this category was named by Nilsen from the perspective of implementation science. In our case, an "implementation theory" might become a Classic theory, and I would like to substitute "health informatics theories" for this 4th category, i.e. theories arising from within the health informatics discipline. An example of an health informatics theory is van der Lei's suggestion that data collected for one clinical purpose can rarely be used for another purpose without careful reassessment [4]

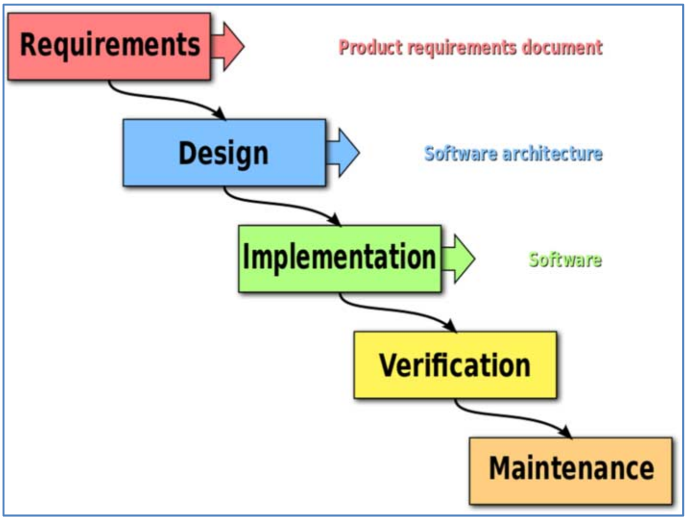


Figure 1. The Waterfall model for software engineering (By Peter Kemp / Paul Smith - Adapted from Paul Smith's work at wikipedia, CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=10633070Carver & Scheier 1982).

Box 1. Schneiderman's checklist of eight user interface feature associated with high usability [2]

1 Strive for consistency
2 Enable frequent users to use shortcuts
3 Offer informative feedback
4 Design dialog to yield closure
5 Offer simple error handling
6 Permit easy reversal of actions
7 Support internal locus of control
8 Reduce short-term memory load

5. Evaluation frameworks: a framework or checklist that specifies aspects of a project that can be evaluated to determine if it is successful. An influential example in health informatics is Kidholm's Model for ASsessment of Telemedicine applications, the MAST framework [5].

In this book, process models and evaluation frameworks are considered to be of less importance than determinant frameworks, classic theories and health informatics theories, so I will confine our discussion to these three categories, referring to them henceforth as "theories".

It is no exaggeration to state that the editors and most authors of chapters in this book would agree that the identification, testing and use of theories is crucial to the future maturation of health informatics as a recognized profession, for at least five reasons. First, no one would argue that we currently know how to produce usable, effective clinical systems every time – indeed, it seems that sometimes success in clinical informatics is the exception rather than the rule [6]. So, we need predictive theories to make health information systems better: that is, more usable, better accepted, more accurate, more clinically and cost effective, and readily transferable to other settings. Second, we need theories to help us build an evidence and scientific basis for our discipline, to help it evolve from a craft - based on anecdote, apprenticeship and learning from mistakes - to a professional engineering discipline [7] similar to, for example, the development of aeronautical engineering. Box 2 describes an example from aeronautical engineering of how formulating and testing a theory became a key method both to enhance aircraft safety and to promote the emergence of a professional discipline.

Third, we need theories (and an understanding about which theory to use, and when) to teach our students and practitioners. Fourth, we need theories to guide organisations procuring systems, so that they can distinguish between theory-based systems that are likely to be effective from atheoretical systems which are less likely to help. Finally, we need a list of tested theories (both useful and useless theories) to help decide rationally whether to carry out a full evaluation of a clinical system following an update or not, according to whether the components that were theory-based are still included. There is an analogy here with medical devices regulation [9]: a previously approved cardiac catheter does not need further testing and regulatory approval if the changes are minor, but it does if the changes are "material". In our case, we could be confident if a lifestyle app, for example, is altered in a minor way, but not if theory-based behavior change features are removed.

Box 2. The importance of theory in aeronautical engineering: a sobering example

In the 1950s there was a disastrous series of 13 aircraft accidents in which the world's first passenger jet airplane, the de Havilland Comet, exploded in midair or on the runway, with a total loss of 426 lives [8]. After some delays, the cause of most of these Comet accidents was traced to cracks in the fuselage of planes, which by then were a few years old. Materials scientists hypothesised that the cracks were caused by metal fatigue starting at the corners of the square fuselage windows and spreading during multiple pressurization cycles. This clear, generic, testable and enduring theory was confirmed by metallurgists examining fragments from crashed planes and subjecting new fuselage segments to multiple compression cycles in a water bath. Fuselage metal fatigue as a cause of aircraft failure is now eliminated by rounded window corners. This is a compelling example of how a theory was identified, tested, then applied universally to make a very complex device more reliable. A long series of such events has led to aeronautical engineering becoming a theory-based discipline in a way that health informatics sadly cannot yet claim to be.

2. Another perspective on "theory", where do relevant theories originate, and which theories are useful?

Although Nilsen defines theory as a set of analytical principles designed to structure our observations, understanding and explanation of a phenomenon in the world, I would advance a slightly nuanced perspective on what definition is needed for "theory" in health informatics: "*a concise, testable predictive principle that can guide the design, development or implementation of clinical information systems*".

Such predictive principles can be derived from an existing theory or can be the basis for a new theory after sufficient testing in multiple settings. Predictive principles derived from theories originate in many disciplines, including psychology; management, implementation or computer science; or healthcare [10]. Some examples:

- Psychology: theories of information design, behavior change or self-efficacy
- Management science: innovation theory, organization theory, marketing theory
- Implementation science: active implementation of guideline recommendations, informed by a study to elicit relevant barriers and enablers, is more effective than simple dissemination
- Computer science: software engineering theories, human computer interaction theories, persuasive technology theory
- Healthcare: investigation is more efficient when test-treatment threshold is considered; prescriptions are safer when drug allergies, interactions and disordered drug metabolism or excretion are considered.

Once we know where we can locate potentially useful theories, we need to understand which predictive principles derived from these existing theories are most likely to be relevant to health informatics, and to be useful so that we can select the most promising. Or we can define and test new predictive principles to help develop new theories. My theory (!) is that, to help our discipline, a predictive principle needs to be clear, predictive, testable, generic but relevant and enduring [10]. Table 1 explains these terms, gives examples and some counter examples of hypothetical predictive principles that would violate each criterion.

Table 1. *Characteristics of a useful predictive principle, examples and counter examples*

Characteristic	Explanation	Example predictive principle	Counter example [explanation]
Clear	Worded so that the implications for system design, development or implementation are explicit	Clinical systems that are problem focused, usable, and incorporate relevant technical standards will be well accepted by users	High quality systems work best [too vague and imprecise to be useful]
Predictive	Applying a predictive theory correctly will result in a clinical information system that is superior in terms of usability, acceptability, effectiveness or cost effectiveness	Clinical information systems that apply Schneiderman's user interface principles will be better accepted by professional and public users	The pricing of clinical information systems depends on the amount of business benefit they help organisations to realize [not predictive]
Testable	Can be readily tested for its relevance in predicting the usability, value or cost effectiveness of health informatics or clinical information systems	User interface designs for clinical information systems that ignore portrait and landscape screen formats will be rejected by tablet computer users	When people believe in a system, it will help them. [Not testable – and a self-fulfilling prophecy]
Relevant	Can be applied to the design, development or implementation of clinical information or clinical research systems	Incorporating Michie's behavior change taxonomy into the design of digital tools to influence health-related behaviours will make them more effective	The development of infection control measures needs to focus on the source and vector of the infection and the nature of the pathogen [not relevant to design of clinical information systems]
Generic	Applicable across wide range of technologies, use cases, users, care settings and health systems	Attention to usability and the balance of perceived benefits and costs, including time to use and privacy risks, will improve the engagement of health professionals with digital tools	Use no more than 15 Cyrillic characters per prompt on ePrescribing app screens for use in Crete [too specific to be useful except in rare contexts]
Enduring	Not likely to be rapidly outmoded by changes in technology or clinical practice	Data that is captured once in a neutral context and is accompanied by meta data is more likely to be reusable	Avoid using batteries with less than 8 hours life in a wearable [will be obsolete once kinetic energy harvesting and thermoelectric generator technologies mature]

3. Testing the validity of a theory and link to evidence based health informatics

One challenge with this new view on theory in health informatics as a predictive principle is that any self-appointed expert can formulate an apparently credible predictive principle that seems to comply with the five criteria listed above and then market it using a catchy acronym (eg. Include Technology When One Risks Knowledge Shrinkage, ITWORKS©), resulting in a generation of clinical information systems that respect the new principle but are actually less usable and effective than those which ignore it. However, the fundamentals of evidence based informatics (EBHI) teach us that expert opinion and authority are not sufficient to provide valid theories, and that principles should be tested before dissemination. So, each predictive principle should be subjected to rigorous evaluation studies in relevant settings, to test if it really does contribute to making systems better. Only if it passes these tests should it be accepted as a predictive principle for our discipline, to be applied in systems development and taught as part of accredited educational programmes for newcomers [10].

The details of designing and carrying out such theory-based evaluations are beyond the scope of this chapter but are summarized in a chapter introducing the concept of evidence-based health informatics [10] and detailed in a textbook on evaluation methods [11]. These evaluation principles need to be understood not only by academics developing and testing new principles but also by system designers and developers applying new principles, so that they can confidently carry out a critical appraisal of the studies that have been conducted to test the principle before applying it. Thus, the central idea discussed in this book of using theories or predictive principles more widely is closely linked with the idea of developing and testing theories central to evidence based health informatics.

Figure 2 below illustrates the suggested process, from identifying a theory relevant to HI, deriving a predictive principle from it then incorporating the principle into system development and testing if this improves the system, for example by making the system more usable, accurate or effective.

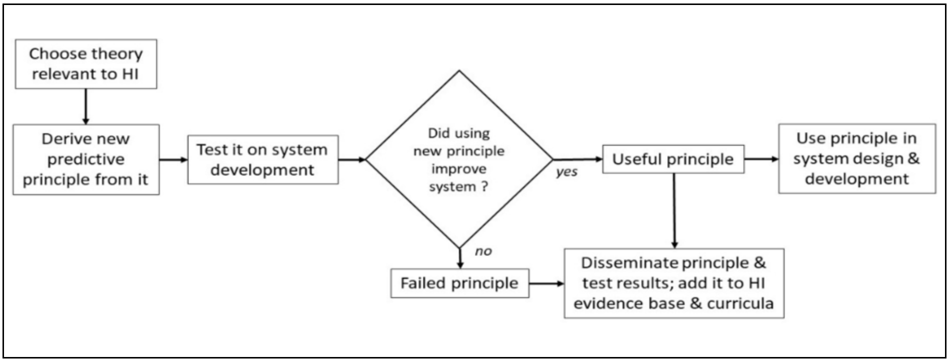


Figure 2. *How to identify and use a useful health informatics principle*

An example of this kind of evaluation is the study we carried out to test the applicability of Fogg's theory of credible website design, based on the design of eCommerce websites, to inform health-related decisions [12]. We designed two version of a website about organ donation with near identical content and usability, one of which followed all Fogg's credibility recommendations while the other site lacked all of these. We then recruited over 800 students via email and randomized them to experience either the credible website or the less credible version. After 4 weeks, we asked participants to join the NHS Organ Transplant Register. Surprisingly, an identical proportion of about 38% joined the register in each group, demonstrating that credible website design had no role in taking this decision [12].

4. Discussion and conclusions

The remaining chapters of this book describe many examples of theories or principles which have been tested in health informatics, with mixed results. However, for the reasons identified in section 1 above, we need to accelerate our progress on theory-based informatics, which requires three specific actions. First, we need to identify more theories that seem relevant to our work from the many existing sources, and derive predictive principles from them for testing. We should also not hesitate to formulate our own testable, generic principles (perhaps to explain failures in a clinical information system we developed, by analogy with the aeronautical engineers investigating the Comet disaster in the 1950s, described above in box 2). Second, we need to test the applicability of each principle in a variety of contexts, to build confidence that the principle – and the theory from which it was derived - does indeed lead to more effective systems [10]. Finally, whether the result of the testing process is positive or negative, we need to work with research and professional organizations at the national and international scale to share that principle and the test results with students and system developers, to encourage them to adopt useful, relevant principles and to drop any that testing shows to be unhelpful, or even harmful. Only this way, in my view, can we move our discipline out of the shadows of authoritarian tradition, superstition or even mysticism, where systems are as likely to harm as to help [13, 14], into the bright light of professionalism where robust, scientifically tested theories and principles guide our work, resulting in predictably usable, safe and effective information systems [15].

Teaching questions for reflection

1. How can the use of tested theories and principles move the health informatics discipline forward as a scientific discipline?
2. What are the potential risks and downsides of a greater reliance on theory?
3. Why are theories or principles advocated by experts not necessarily useful to guide the development of better clinical information systems?
4. How would you test the impact of a new principle that claims to guide the design of safer ePrescribing systems?
5. Will we ever have a Grand Theory of health informatics? If so, could it pass all five criteria for a useful predictive principle listed in the table in section 2?

References

[1] P. Nilsen. Making sense of implementation theories, models and frameworks. *Implement Science* **10** (2015), 53.

[2] User interface design principles. Section 3.3.4 in B. Shneiderman, C. Plaisant, M. Cohen, S. Jacobs, and N. Elmqvist. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*: Sixth Edition, Pearson (May 16).

[3] S. Michie, M.M. van Stralen, R. West. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Science* **6** (2011), 42.

[4] J. van der Lei. Use and abuse of computer-stored medical records. *Methods Inf Med.* **2** (1991), 79-80.

[5] K. Kidholm, A.G. Ekeland, L.K. Jensen, J. Rasmussen, C.D. Pedersen, A. Bowes, S.A. Flottorp, M. Bech. A model for assessment of telemedicine applications: mast. *Int J Technol Assess Health Care.* **1** (2012), 44-51.

[6] R. Koppel, J.P. Metlay, A. Cohen, B. Abaluck, A.R. Localio, S.E. Kimmel, B.L. Strom. Role of computerized physician order entry systems in facilitating medication errors. *JAMA.* **293** (2005), 1197-203.
[7] H. Heathfield, J. Wyatt. The road to professionalism in medical informatics: a proposal for debate. *Methods Inf Med.* **34** (1995), 426-33.
[8] Anon. The de Haviland Comet. https://en.wikipedia.org/wiki/De_Havilland_Comet
[9] European Commission. Guideline on the details of the various categories of variations to the terms of marketing authorisations for medicinal products for human use and veterinary medicinal products. https://ec.europa.eu/health/sites/health/files/files/betterreg/pharmacos/classification_guideline_adopted.pdf
[10] J.C. Wyatt. Evidence-based Health Informatics and the Scientific Development of the Field. *Stud Health Technol Inform*. **222** (2016), 14-24. (Also published as Chapter 2 in: Ammenwerth E, Rigby M. Evidence based health informatics. Amsterdam: IOS Press ISBN 978-1-61499-634-7 (print) | 978-1-61499-635-4 (online)
[11] C. Friedman, J.C. Wyatt. *Evaluation Methods in Biomedical Informatics* (second edition 2006). New York: Springer https://www.springer.com/gb/book/9780387258898
[12] T.J. Nind, I.W. Ricketts, F.F. Sniehotta, J.C. Wyatt, S. Cole. Do credibility cues affect the persuasiveness of a web site? Symposium: Influencing People with Information. Aberdeen. (2012)
[13] E. Ammenwerth, N.T. Shaw. Bad health informatics can kill--is evaluation the answer? *Methods Inf Med*. **44** (2005), 1-3.
[14] Bad health informatics can kill - https://iig.umit.at/efmi/badinformatics.htm
[15] J. Wyatt. Medical informatics, artefacts or science? *Methods Inf Med*. **3** (1996), 197-200

Part 1

Information Science and Technology Theories

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

doi:10.3233/SHTI190107

General System Theory and the Use of Process Mining to Improve Care Pathways

Owen JOHNSON [a,1]
[a] *School of Computing, University of Leeds, Leeds, UK*

Abstract. General System Theory was proposed in the post-war period as a unifying framework for interdisciplinary science based on the idea that *systems* have a set of similar properties and characteristics regardless of discipline. General System Theory laid the foundations for talking about things in terms of systems, many of its terms are now embedded in everyday language and it underpins a broad range of *systems approaches* and *systems thinking*. This chapter will describe the key elements of the original General System Theory (GST) including *control*, *feedback*, *emergence*, *holism* and the notion of a *hierarchy* of systems within systems. It will review the origin, content and foundational role of systems theory in biology, medicine, computer science, organizational theory and its central contribution to health informatics. In recent years, healthcare organizations have been encouraged to see themselves within the context of learning health systems (LHS) and to use emerging big data analytics techniques such as process mining to develop better, integrated and personalized pathways of care for patients. We use GST to reflect on these emerging approaches through a discussion and case study on recent work in urgent and emergency care. Our aim is to trace the influence of GST through emerging LHS ideas and use the framework of GST to reflect on the opportunities and limitations of our process mining approach. In particular, we will reflect on how GST can explain successes and failure in the application of process mining to care pathways and the challenges and opportunities ahead.

Keywords. General System Theory, Learning Health Systems, Process Mining

Learning objectives

After reading this chapter, the reader will be able to:

1. Review general system theory and the rich set of perspectives it brings to the understanding of health informatics in modern organizations.
2. Illustrate the application of general system theory to current challenges in healthcare.
3. Use general system theory as a framework to review data driven approaches to care pathway improvement with a specific focus on process mining.
4. Use general system theory as a perspective to reflect on the opportunities for learning health systems that focus on care pathway improvement.

[1] Corresponding Author: Owen Johnson, E-mail: o.a.johnson@leeds.ac.uk

1. Introduction to General System Theory

1.1. The origins of General System Theory

Systems approaches to thinking about the world run through much of Western philosophical thought. Eastern traditions have similarly emphasized systems concepts such as holism and the balance between change and homeostasis [1]. Our modern understanding of systems can be traced to General System Theory, proposed by Karl Ludwig von Bertalanffy [2] as a unifying framework for systems that is equally applicable to organisms and organizations. Branches of systems theory underpin software engineering, soft systems methods, cybernetics and Artificial Intelligence (AI). Health informaticians should see *systems theory* as a fundamental and powerful tool in their professional role and no textbook reviewing interdisciplinary theories for health informatics would be complete without a discussion on General System Theory and the impact that modern systems approaches have had on both healthcare and informatics.

The *language of systems* permeates all aspects of computing, information technology and the computer systems that we health informatics practitioners design, implement and study. Our computer systems are a special case of more general systems. They are different from, but also similar to many other types of systems, and of course, they are a key component in what are increasingly being called *healthcare systems* – that complex set of organizations and relationships that provide healthcare to large populations. A systems approach should be particularly appealing to health informaticians because systems perspectives and principles are applicable in medicine, biomedical sciences, systemic approaches to therapy, informatics systems and the organization of healthcare services [3]. New health informaticians may be surprised at the extent to which *systems* are found in medicine and biology and the importance of *systems thinking* in the understanding of the human body, its healthy maintenance and its responses to diseases and therapies.

Both computers and people are complex systems. In the middle ground between computer systems and a medic's understanding of biologic systems lies the myriad web of healthcare organizations, processes, care pathways and health delivery systems which health informatics seeks to improve. In the complex space of healthcare the words "systems" and "systems approaches" are often used rather carelessly and with little understanding or awareness of the *science of systems*. In this chapter, we aim to acquaint the health informatics practitioner with the theoretical base in *systems* that underpin both medicine and computer science and make the case for leveraging General Systems Theory as a toolset for addressing applied healthcare challenges.

General System Theory (GST from here onwards) was developed by a biologist, Karl Ludwig von Bertalanffy (1901-1972), was given support and a framework [4] by an economist, Kenneth Boulding (1910-1993) and has been subsequently refined and developed by many other scientists from a diverse range of disciplines. Bertalanffy developed his ideas for GST before and during the Second World War but did not publish them until afterwards and at a time where there was an explosion of post-war systems ideas and approaches. This interest in systems coalesced into a wider systems movement and included developing the principles and theories for the first computer-based systems. The foundational role of GST was that it provided this new systems movement with the belief that there was unifying framework underpinning their efforts. GST made the case for a single language for systems and for systems approaches as science [2].

GST was a development of Bertalanffy's work on *open systems* in biology. In physics, the laws of thermodynamics are based on conservation of energy and a tendency towards entropy (disorder) in a theoretical "isolated system". Bertalanffy noted that such isolated (or closed) systems rarely, if ever, exist in nature and, in biology, organismic systems (his phrase) tend towards order rather than disorder and, most obviously, organismic systems can grow and replicate as they interact with their environment, exchanging energy, matter and information. Many organismic systems are able to dynamically respond to their environment in order to maintain a steady state (homeostasis) in, for example, body temperature. More generally, the tendency towards order can be found in atoms, molecules, cells, organs, organisms such as people and organizations such as social groups and even health care providers. Structures emerge based on finding effective relationships between components whether these are protons and neutrons in an atom or a surgical team trying to save the life of a critically ill patient. Bertalanffy's development of GST was motivated by his desire to provide a fundamental language of systems that would improve scientific understanding across all disciplines [5].

1.2. What is General System Theory?

General System Theory in the narrowest sense was defined by Bertalanffy as the attempt to derive a general definition of "system" as a **complex of interacting components that together have the characteristics of an organized whole** [5, pg 91]. The emphasis of a system as an "organized whole" incorporates the concept of *holism* developed by Aristotle and commonly expressed as "the whole is more than the sum of its parts". GST makes the connection that holism is an *emergent* characteristic of systems as a product of the relationships between its components as they work together to collectively interact with their environment.

There are three key principles that follow from this general definition and the emphasis on holism. Firstly, GST asserts that this definition of a system should be generally applicable across all disciplines and that the systems perspective can generate new, and hopefully useful, insights.

Secondly, GST states that components of systems are often systems in their own right. Each member of a surgical team is also a person with similar but also unique emergent characteristics that might included their degree of experience in the specific role, their skills but also their affinity with other members of the team and degree of tiredness, hunger etc. which could be traced to their digestive systems and maintenance of blood sugar levels. From the perspective of GST, systems can be seen as being both composed of, and existing within, a *hierarchy* of systems. Our surgical team may be part of a busy Accident and Emergency department within a large hospital that is part of a larger healthcare provider and a regional or national health system. The team's performance will be affected by their immediate environment, which will include other systems (teams, departments, etc) within the hospital that it interacts with (in collaboration with or even in competition against) and also external environment factors such as the arrival of more patients.

Thirdly, GST places the emphasis on the *relationship* between components rather than simply the components themselves. The fact that surgical teams generally cope so well with all the complexity thrown at them is a testament to the relationship between team members - roles are clear but also sufficiently flexible and dynamic to adapt quickly to each other's needs as well as the patient's. An emergent property of a surgical team is that it is good at doing the appropriate medical or surgical interventions. The same people given the right training and tasked with organizing the introduction of a new computer system might struggle to work as an effective team simply because the relationships required are likely to be very different.

Two major criticisms of General Systems Theory are worth reflecting on at this point – one is that it is too general, and the other is that it is not really a theory. Bertalanffy was keen to insist that the aim of GST was not to provide a general theory of everything that would be so general as to have no practical application [2]. In his view, GST should provide a perspective *where it is useful* in providing a language or a framework for thinking about and discussing systems, particularly between disciplines that could benefit from sharing fresh ideas. While there are a dizzying range of potential hierarchies of interacting systems and sub systems in our surgical team example, a sensible use of GST is to focus on just those systems where a systems perspective generates fresh and useful insight. The second criticism of GST as "not really a theory" has some foundation. Bertalanffy himself argued that GST was conceived as a working hypothesis, a goal rather than a clear axiom [5]. Tom Mandel in "Yes, there is a general system principle, No it is not a theory" [6] makes a fair case for GST being regarded as a principle although the counter argument might be that the theory is that the principle applies. Semantics aside, it is perhaps best to regard GST, as Bertalanffy intended, as a "theoretical model" whose value lies in the practical "explanation, prediction and control of hitherto unexplored phenomenon" [5, pg 99].

1.3. Extensions to General System Theory

In Advances in General System Theory [5], Bertalanffy explored how the explosion of post-war systems approaches might fit with GST to provide a broader general theory of systems developing the principles of *communication* and *control* that describe how systems work. *Shannon's Information Theory* introduced the concept of information as quantity and "negative entropy" (information reduces uncertainty) and developed the principles for describing information transmission used in computer science[2]. Systems use information from their environment to reduce uncertainty about the range of appropriate responses, for example, when a medic uses diagnostic results to rule out possible diseases, narrowing down the options to identify the most likely disease and decide on the best treatment. *Cybernetics*, based on the role of information feedback in circular causal chains, helps explain how systems can be self-controlling. As early as 1948, William Ross Ashby applied cybernetic principles to build a synthetic brain, called the Homeostat, from four interlinked air force bomb control units that worked together as a system to maintain *homeostasis* through reinforcement and learning. Ross Ashby's *Law of Requisite Variety* is useful here; the *survival* of a system over time depends on it retaining sufficient (requisite) variety in its internal structure to respond to the variety in its environment; systems fail when they are unable to adapt to their environment. *Game Theory* describes logical decision making in humans, animals, and computers and provides insights into how some systems are maintained through *competition* between components where each component competes to maximize gain and minimize loss. In organizational systems, market forces often dominate - students compete for higher marks, professionals compete for salary, roles and kudos, and both private and public organizations compete for work and resources. GST includes the idea that relationships

[2] Discussed further in Chapter 3, "Information theory and medical decision making".

between system components can be competitive; in many systems, it is the dynamic tension of relationships between components that creates structures that stand the test of time. Similarly, Bertalanffy argues that GST also embraces *decision theory*, which analyses rational choices within human organizations, and *network* and *graph theory*, which can help GST develop models of the complex relations between system components in, for example, social networks. GST expanded from a theory into an ambitious project to join together disparate systems related ideas.

Perhaps the best attempt to provide a useful synthesis of all the multiple systems theory perspectives comes from Ken Boulding's (1956) paper titled "General systems theory: The skeleton of science"[4]. Boulding's framework categorizes various types of system in terms of eight levels of increasing sophistication that could be seen as systems archetypes. Level 1 (Simple Structure) are borderline candidates for systems in that they have physical structure but are essentially static, for example a rock. In healthcare we might think of objects such as a scalpel, a bed or a room, such objects still have an emergent property of wholeness and, for human created artifacts, often some discernible purpose. Level 2 (Clockwork) are more sophisticated than Level 1 in that they have movement and may maintain an equilibrium but such movement is predetermined, most obviously a clock-work clock and other simple machines but also the solar system. In healthcare, such concepts underpin stochastic dynamic modeling of, for example, the seasonal rise and fall of demand. Level 3 (Control Mechanisms) are Level 2 systems that also have some element of *information* closed-loop control, the classic example being a thermostat which turns heating on or off based on comparing the feedback of the current temperature to the control setting[3]. These are the principles of *cybernetics* in computing and *homeostasis* in biology and in management underlie principles of stock control and resource planning now often encoded within enterprise resource management systems. Level 4 (Open Systems) are Level 3 systems that have a self-maintaining structure in constant interaction with its environment, such a definition might include a flame or a river but more generally is the essence of simple life, a cell or a virus where we can add in the property of being able to self-reproduce. Level 5 (Plant) are Level 4 systems which have an organized whole based on a structure of differentiated and mutually dependent parts, for example plants where roots, leaves, seeds etc are functional parts themselves composed of specialist cells (i.e. Level 4 systems). Level 6 (Animal) are Level 5 systems which display *intelligence*, typically with sophisticated information intake, processing and control including the construction of a *knowledge* structure that enables them to compete (*Game Theory*) and make informed decisions (*Decision Theory*). Level 7 (Human) is distinguished by adding *self-consciousness* and, one would hope, more intelligence, greater reasoning based on knowledge and a capacity for more complex processing of symbols such as in the use of language. Level 8 (Social Organization) are the complex collections of people in various *roles* that manifest as discernible systems. An individual may simultaneously be a mother (and a daughter) in a family, a surgeon in a surgical team, an employee within a healthcare organization and a researcher doing a part time PhD at a university. The family, surgical team, healthcare provider and university all fit the definition of Level 8 systems and are *social networks* of people. Experience tells us that all of these can be hugely complicated, constantly changing and yet somehow their structures persist and evolve through changing relationships and the arrival and departure of new people. In Boulding's words, Level 8 includes "human life and society in all its complexity and richness" [4, pg 200].

[3] See also Chapter 14, "Control Theory to design and evaluate audit and feedback interventions".

Boulding's levels should not be mistaken for an attempt to provide a definitive taxonomy of GST or of life. It does however provide, and is best used as, a simple framework for discussing system models of increasing complexity. GST also provides a starting point for the rich world of systems thinking and systems approaches that can help health informatics practitioner understanding and improve the use of health informatics in modern organizations. Such approaches include systems engineering, Peter Checkland's Soft Systems Method [1], complexity science, systems dynamics, simulation and Peter Senge's Learning Organizations [7].

2. Using GST in Health Informatics

2.1. How health informatics professionals can use GST

Health informatics professionals can:

1) Use GST in its narrowest sense to identify, model and define a system of interest following the definition of GST in Section 1.2. A careful choice of boundary is essential as the components and relationships within the system should be directly responsible for the system appearing as a coherent whole. GST forces deep reflection on what the system actually is, how it survives over time, its structure and environment. A good understanding of how and why a surgical team works well should be an essential pre-requisite to an implementation project introducing a new informatics solution that is expected to help their performance. Conversely, of course, it can help understand why health informatics projects often fail. We would encourage the former.

2) Use the language of GST for interdisciplinary communication. We have italicized most of the key GST terms in this chapter and the informatics practitioner who is familiar with and can use these terms in discussion with healthcare professionals (and even managers) should find that they are speaking a common language if only because most will have learnt them in biology classes.

3) Develop their understanding of GST into a broader systems approach to problem solving. There are many good books, courses and online material that are linked to and build on GST and systems approaches. Once you have started thinking in systems, it is difficult to stop and there are many practitioners who consider systems thinking has transformed their professional approach.

The following examples of the applications of GST within health informatics will, we hope, illustrate the scope and potential.

2.2. Applications of GST in Healthcare Computing

Our modern computer systems were first developed within the climate of the post-war systems movement and computer science has contributed to, and benefited from, GST. In common with other systems, computer systems have components (software and hardware) and relationships (interfaces, dependencies and networks) and we can describe these in terms of *inputs*, *processes*, *outputs*, feedback and control. Component based and layered architectures are designed to *manage complexity* while delivering functionality and performance at scale. Most people know from experience that some computer systems are better than others and that some can crash or slow down unexpectedly. Computer system performance (and usability, security and other non-functional characteristics) are emergent properties of the system-as-a-whole. The complexity of modern systems is such that solving one performance issue or bug may introduce others and a holistic perspective on the system together with a deep respect for the complexity of its internal structure becomes essential.

As our computer systems have become more complex, they have become, following Boulding's Level 7, more human. Holistically they can display emergent properties of being buggy, annoying, slow stubborn, inflexible – to the extent that we may find ourselves shouting "stupid computer" at them or complaining about them as though they were a troublesome colleague. From a GST perspective, none of this should be a surprise - most health informatics systems fit comfortably into Boulding's definitions of Level 5 and above and may have many of the characteristics of Level 7, and perhaps Level 8 too. Especially as modern advances in computing such as AI, neural networks, distributed systems and edge computing increasingly follow biologic models of systems of competing sub-components. The result is that even their designers cannot know exactly how they work. For healthcare this presents an unusual problem: should clinicians trust a computer system that no-one can adequately explain? Medical devices have been regulated on the basis that their programming is rules-based (GST Level 3 and 4) but complexity in general and medical AI in particular have advanced computing well beyond these levels. GST may be needed to help regulation, legislation, the professions and society adjust to human-level computer-based systems.

One significant difference between all computer systems and all biologic systems is the relationship with data. Biologic systems process and act on information and store useful information and successful responses to it as knowledge for future reference, and they have used this learning system process to evolve successful survival skills over many thousands of years. Our current computing systems are an awkward fit with GST; they are less than 70 years old and they work differently. Specifically, they can and do store huge amounts of raw data and it is their reliance on data, rather than information and knowledge that can make them appear "stupid". Future, bio-inspired computing may evolve similar intelligence but for now the key opportunity for organizations is to mine the wealth of *big data* stored within legacy computer systems. In healthcare, data mining of electronic health records is seen as having the potential to transform our understanding of medicine [8]. Locked away in these records is the history of millions of clinical encounters and their successful or unsuccessful outcomes.

2.3. Applications of GST in Learning Health Systems

In health informatics, there has been growing interest in *Learning Health Systems*, a phrase coined by Charles Friedman [9] in the USA which envisaged rapid learning based on a federated, national approach to exploiting EHR data gathered by different US healthcare providers. More generally, Learning Health Systems are seen as organization-wide or pan-organizational regional and national systems that deliver healthcare to a large population. In Friedman's vision there is a symbiotic relationship between the health provider system and the health information systems that it uses. The *Learning Organization* concept was developed from systems theory by management theorists, notably Peter Senge [7]. In learning organizations, systems approaches that reward effective learning are embedded within management culture at all levels of hierarchy. The organization is seen as organic with structures evolving through continuous learning to meet changing environments and ensure survival in a fast paced, ever changing world.

Exactly as described in GST. In Learning Health Systems (LHS), these ideas are extended to include developing new medical learning and there is a strong emphasis on the use of health informatics solutions as both the provider of the data that will be used for *evidence-based medicine* and the vehicle for delivering knowledge to the clinical teams through automated decision support and workflow management.

The Heimdall Framework [10] provides a taxonomy of types of learning health system where new clinical insight and patient process improvements are driven by the analysis of data from the electronic health record (EHR) and other health information systems. In GST terms, clinical and management control is informed by feedback about processes and outcomes and is implemented as interventions to the inputs and process. More data, faster data flows and improved analytical abilities improve control and the organization's long-term ability to continuously learn and adapt to its changing environment. A key insight from GST is that of systems-within-systems, each contributing to overall success. An LHS approach can therefore be applied to a surgical team, a ward, a department or clinical specialty as well as the organizational, regional and national systems in Friedman's vision. Following GST carefully would suggest that LHS should indeed be implemented at all levels of the organizational hierarchy including the individual human as reflective practitioner. Adoption of integrated informatics solutions, interoperability standards and improved methods for mining health data are essential for LHS but the long term vision is of systems that self-learn through embedded AI and a new generation of digital-native clinicians who are part of, but remain firmly in control of, their health system. LHS is seen as a driver for health informatics but to succeed it requires the deeper understanding of the relationships between organizational structure, people, processes and technology that comes from applying GST.

2.4. Applications of GST in Process Mining of Care Pathways

The care pathway is a commonly used concept for considering how the processes of delivering healthcare should best be organized around the needs of the patient [11]. A care pathway is a design template for a healthcare process – it describes the sequence of care that is recommended for patients with similar conditions requiring similar treatment. Comparing the actual care that patients received as recorded in the EHR against the intended care pathway should help healthcare organizations understand the gap between what they think they are doing and what they are actually doing, a key requirement for learning. Coiera [12] suggests that LHS should use process mining to develop automated process-level metrics and identify common multi-variate process patterns to help better understand how healthcare delivery is structured. Process mining is a set of big data analytics tools and techniques that use time-series event data to specifically address process characteristics and there is growing interest in process mining in healthcare [13]. Ronnie Mans and Wil van der Aalst [14] provide a comprehensive guide to process mining in healthcare including health reference models and pathways. Process mining has been combined with process simulation to create a mixed methods approach to support the development of LHS [11]. In the following example we illustrate how process mining of a care pathway fits with GST and an LHS vision.

3. Success factors in process mining of care pathways

3.1. Connected Health Cities

The Connected Health Cities (CHC) project in the North of England aims to implement a region-wide LHS through a range of initiatives linking and using health data and sharing insights and best practice (www.connectedhealthcities.org). The approach has included the development of federated data repositories of EHR data as advocated by Friedman, the development of a learning culture for sharing and disseminating knowledge and a focus on care pathways that can be mined, analyzed and improved. Challenges have included: developing architectures and consent models for ethical access to health data; linkage of health data from different sources, standards and variable data quality; engagement with multi-disciplinary teams across multiple organizations; engagement with busy clinicians and already stressed organizations; and the development of better methods for process mining of care pathways. Solutions have included: national level engagement on legal and ethical frameworks; public engagement through a social media campaign (called #datasaveslives) and citizens juries; Trusted Research Environments (TREs) for the secure curation of data; developing experience in multi-disciplinary collaboration; a focus on specific high-impact problem areas; and ClearPath, a novel method for care pathway process analysis that draws on GST and, more generally from a systems thinking approach.

3.2. The ClearPath Method

The ClearPath method [11] is an extension of an established process mining method (called PM^2, see [14]) that incorporates a stronger systems method of enquiry and produces care pathway simulations that can be used for experimentation and learning. In our work in this area it became evident that a more holistic systems approach was essential to address what have been called "data quality" issues. From the perspective of GST we see health data not as the product of a machine but as the product of a highly complex sociotechnical healthcare system that is evolving, adapting and responding to its environment. We would argue that the failure of "big data" methods in healthcare is due to a failure to apply GST. A conventional approach to healthcare data mining includes complaining about data quality, cleaning data to suit the analysis and assuming that more data means less unknown systemic bias. The reality of healthcare data is that it is messy and incomplete, it can shed some light on the activity of busy clinicians and the administration of healthcare processes but with different systems used differently by different departments, highly variable pathways and moving systems boundaries the only real certainty is that data will be different between systems and over time. Recent advances in process mining recognize this phenomenon as process evolution or "concept drift" and new techniques such as applying sliding time windows to spot changes in process are being developed with some success [15].

Our approach within the CHC project has been to combine process mining of EHR data with a systems approach to enquiry. Following GST, the starting point is to identify and define a system of study that has a clear boundary and a single clear structure. For example we have worked with a number of urgent care departments and have treated each one as a separate discrete system, resisting the temptation to aggregate urgent care data across the region because such an aggregated view would fail GST's test of what is a system, a common mistake made by those who advocate big data in healthcare. We have however modelled urgent care as a part of the larger system of a hospital and the wider health system, for example across a district and a city region, using a systems-within-systems approach that does fit well with GST. We recognize health systems as GST Level 8 open systems; the relationship between system and environment is complex and evolving. In this context, process mining is useful in looking for those patterns and structures that emerge from a holistic view of the system.

In the UK there have been national targets for at least 95% of patients attending Accident and Emergency departments to be admitted, transferred or discharged within four hours. A pattern that emerges from process mining many such departments is that a median of 3.9 hours is common. Root cause analysis discussions with domain experts suggests this is game theory at work. The national target leads to the perverse behavior that the staff wait until, and then respond to, the impending deadline perhaps also believing that a full waiting room and a long wait will discourage less seriously ill patients. We also found evidence that the patterns and sequences of processes change during the day. Standard process conformance metrics were noticeably at their worst around early evening when routine processes give way to a period of apparent chaos with, for example, beds being requested for patients that had not yet been seen by a clinical specialist. We traced this flurry of activity to the time when the overnight shift starts work and a new allocation of beds become available; our discussions suggest that the new shift prioritize operational concerns such as booking beds over the routine updating of the computer system. In both cases these are very human activity patterns that can be explained through GST and only revealed by systematic enquiry.

The other contribution from GST has been the construction of models to represent systems of study. Simple models such as process maps and mathematical formulae can be seen as GST Level 2 or 3 and therefore inadequate for explaining the behavior of a GST Level 8 organization. In the ClearPath method we use a care pathway simulation tool called NETIMIS (www.netimis.com) to present dynamic, runnable models back to multidisciplinary teams as part of a facilitated discussion about care pathway improvement. Simulation modelling might be seen as GST Level 4 and therefore inadequate in capturing the complexity of real-life healthcare. However, the real learning in LHS is still done by people so the discussion and the interactions and ideas it sparks are the real outputs of process mining of care pathways.

4. Discussion

4.1. Is GST relevant to modern health informatics?

The enduring strength of GST is that it opens a window into a powerful way of viewing the world. At its most general, it sees the world as made of systems many of which are dynamic, complex and ever changing _ a melting pot of complexity where structures still emerge and have permanence while the relationships that hold them together are maintained, a wave crashing on a beach, a flight of birds forming a characteristic “V” shape. In our healthcare contexts, a patient’s body fighting serious infection and a surgical team at the end of a tough shift while also perhaps battling with a stubborn computer system. Or the cash-strapped health provider organization that spent too much procuring that computer system because it lacked the internal competencies to appreciate the importance of health informatics.

One student on a recent Systems Thinking course said they found GST difficult because “anything could be seen as a system depending on the boundaries you set”. The student was in one sense correct, but in GST we also expect systems to have emergent structures and simplicity. We recognize that both a wave and a hospital are actually very complex but are happy to accept they exist as systems that we can observe and reason about. Choosing boundaries wisely is important.

The challenges facing health informatics professionals are getting harder not simpler. Many healthcare organizations have successfully implemented health information systems and are now asking how they can use their computer systems to improve their internal structures, processes and deliver better care. We would recommend GST and a systems approach to help make a hard job somewhat easier and more rewarding.

Teaching questions for reflection

1. Reflect on a health informatics system that you are familiar with; write a definition of the system following the definition of GST in Section 1.2. Describe the system in terms of its most significant components and their relationships. Reflecting on the healthcare environment where this informatics system is used, identify a system of healthcare provision and write a similar definition and description. How should the health informatics system contribute to the “survival” (continued viability and effective working) of the healthcare system it is part of?
2. Discuss your understanding of General Systems Theory with people from a range of disciplines (clinical, informatics, management etc.). Looking through the italicized terms in this chapter ask your colleagues whether they recognize these terms and whether they have the same meaning regardless of discipline.
3. From the perspective of GST, our complex healthcare systems can be seen to be in a state of continuing flux and the data in our health informatics systems reflects this. Can Artificial Intelligence (AI) that has been trained on such highly variable data ever be considered safe for clinical use in these constantly changing environments?
4. Reflect on a care pathway that you are familiar with; how could you help implement a learning health system that used health informatics to capture data that would help health care professionals continually learn about and improve the pathway?

References

[1] P. Checkland, *Systems Thinking, Systems Practice*, Wiley, NY, 1981.

[2] K.L. von Bertalanffy. An outline of general system theory. *British Journal for the Philosophy of Science,* **1**, 1950. 134-165.

[3] E. Coiera, *Guide to Health Informatics,* CRC press, NY, 2015.

[4] K.E. Boulding, General systems theory—the skeleton of science, *Management Science,*1956 Apr;2(3):197-208.

[5] K. L. von Bertalanffy, *General System Theory: Foundations, Development, Applications,* NY, 1968.

[6] T.M. Mandel, Yes, there is a general system principle, No it is not a theory. In Proceedings of the 56th Annual Meeting of the International Society for the Systems Sciences (ISSS) 2012, San Jose, CA, USA 2013 Mar 8.

[7] P. Senge, *The fifth discipline, the art and practice of the learning organization*, Wiley, NY, 1991.

[8] T.B. Murdoch, A.S. Detsky. The inevitable application of big data to health care. *JAMA* 2013;309 (13):1351–2.
[9] C. P. Friedman, A. K. Wong, D. Blumenthal, Achieving a Nationwide Learning Health System. *Science Translational Medicine.* 2, 57cm29 (2010).
[10] S. McLachlan, H.W. Potts, K. Dube, D. Buchanan, S. Lean, T. Gallagher, O.A. Johnson, B. Daley, W. Marsh, N. Fenton. The Heimdall Framework for supporting characterisation of learning health systems. *Journal of Innovation in Health Informatics*, **25:**2 (2018), 77-87.
[11] O.A. Johnson, T. Ba Dhafari, A. Kurniati, F. Fox, E. Rojas. The ClearPath Method for Care Pathway Process Mining and Simulation. In *PODS4H, 16th International Conference on Business Process Management (BPM 2018)*, Sydney, Australia 2018 Sept 10.
[12] E. Coiera. The forgetting health system. *Learning Health Systems*. 2017 Oct;1(4):e10023.
[13] E. Rojas, J. Munoz-Gama, M. Sepúlveda, D. Capurro. Process mining in healthcare: A literature review. *Journal of Biomedical Informatics*. **1:**61 (2016), 224-36.
[14] R.S. Mans, W.M. van der Aalst, R.J. Vanwersch. *Process mining in healthcare: Evaluating and exploiting operational healthcare processes.* Heidelberg: Springer; 2015 Mar 12.
[15] T. Li, T. He, Z. Wang, Y. Zhang, D. Chu. Unraveling Process Evolution by Handling Concept Drifts in Process Mining. In *Services Computing (SCC)*, 2017 IEEE International Conference on 2017 Jun 25 (pp. 442-449).

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

doi:10.3233/SHTI190108

Information Theory and Medical Decision Making

Paul KRAUSE[a,1]
[a] *Department of Computer Science, University of Surrey, United Kingdom*

Abstract. Information theory has gained application in a wide range of disciplines, including statistical inference, natural language processing, cryptography and molecular biology. However, its usage is less pronounced in medical science. In this chapter, we illustrate a number of approaches that have been taken to applying concepts from information theory to enhance medical decision making. We start with an introduction to information theory itself, and the foundational concepts of information content and entropy. We then illustrate how relative entropy can be used to identify the most informative test at a particular stage in a diagnosis. In the case of a binary outcome from a test, Shannon entropy can be used to identify the range of values of test results over which that test provides useful information about the patient's state. This, of course, is not the only method that is available, but it can provide an easily interpretable visualization. The chapter then moves on to introduce the more advanced concepts of conditional entropy and mutual information and shows how these can be used to prioritise and identify redundancies in clinical tests. Finally, we discuss the experience gained so far and conclude that there is value in providing an informed foundation for the broad application of information theory to medical decision making.

Keywords. Shannon entropy; Relative entropy; Conditional entropy; Mutual information; Medical diagnosis

Learning objectives

After reading this chapter, the reader will be able to:

1. Understand the basic concepts of information theory: information content; Shannon entropy; relative entropy.
2. Understand how these concepts can be applied to medical decision making at a general level.
3. Understand how the more advanced concepts, conditional entropy and mutual information, could provide deeper insights into the potential redundancies in laboratory tests.

1. Introduction to Information Theory

Information theory has gained application in a wide range of disciplines, including statistical inference, natural language processing, cryptography and molecular biology. It covers the study of the transmission, processing, extraction, and utilization of information at a foundational, mathematical level. A fundamental goal of information

[1] Corresponding Author: Paul Krause; E-mail: p.krause@surrey.ac.uk

theory is to provide a sound basis for optimising the amount of information that can be extracted from a specific situation. Many of the outcomes from the study of information theory have been reduced to engineering practice in a wide range of disciplines, from Artificial Intelligence and Machine Learning, to cybernetics and complexity science. So, the question naturally arises: could it be used to inform the practice of medicine. This is the topic of the current chapter.

Central to Information Theory is the study of situations where one agent (the *transmitter*) conveys some message over a *channel* to another agent (the *receiver*). This is typically performed by having the transmitter send a series of partial messages. In the case of the Internet, for example, the Transmission Control Protocol (TCP) defines how a message may be broken down into packets before sending, enabling the resulting packets to be reassembled in the correct order by the receiver. Each of these partial messages can be thought of as resolving some measure of uncertainty in the receiver as to the content of the original message. The measure of uncertainty resolved by a partial message is its *information content*.

Let us start with a schematic of a general communication system, redrawn after Shannon's original paper [9].

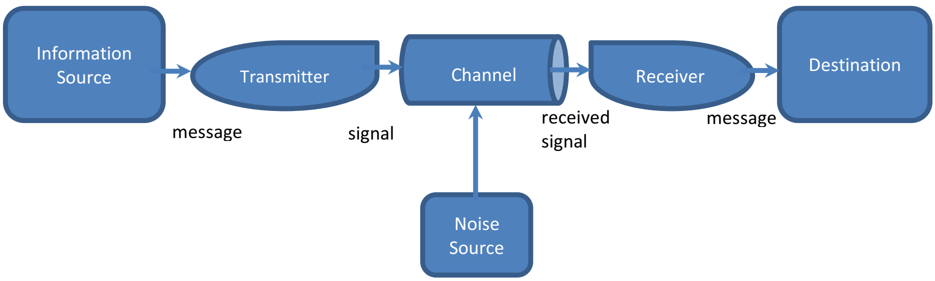


Figure 1. Schematic of a general information system.

We start with an *information source*, which generates a *message* or sequence of messages which are intended to be communicated to a *destination*. The destination is assumed to be remote from the information source. Hence, the message needs to be converted by a *transmitter* into a *signal* that is in a suitable form to be transmitted through some channel, after which the *received signal* is converted back into a suitable format by a *receiver* to enable it to be easily interpreted at the *destination*.

The challenge of communication theory is to understand how information that is transmitted from the source can be completely and correctly received and interpreted by the destination. This is a challenge because in general any communication channel will have an associated *noise source* that may corrupt, to a greater or lesser extent, the transmitted signal before it is received (by altering or even losing components of the signal). Furthermore, we cannot be certain that the transmitter and the receiver are perfect converters of message to signal, and signal to message, respectively.

In this chapter, we will show how viewing diagnosis as embedded within a communication system can lead to an information theoretic perspective on medical diagnosis. Each test or intervention can be seen as a partial message leading towards the desired complete message that provides sufficient information to confirm a diagnosis. At any stage in an investigation, one would then select the next test as the one that would maximise the information gained. The important point about this perspective is that it provides a rational basis for identifying which test to perform at each stage of an engagement with a patient.

It is quite straightforward to map the model in Figure 1 onto the specifics of telegraphy, radio or television broadcasting, and the internet, for example. However, it will be of more interest in the current context if we instantiate the model within a clinical setting.

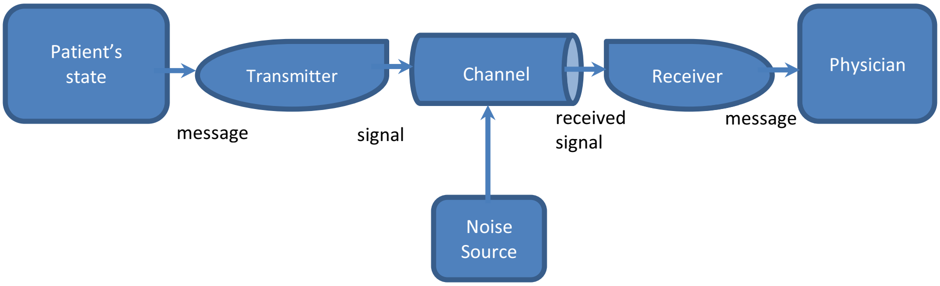


Figure 2. The communication system in a clinical setting.

We have actually changed very little in Figure 2 compared with Figure 1: the information source has become the patient's state; the destination has become the physician who has responsibility for performing the diagnosis of the patient's state.

We do still need to keep the model quite general. The transmitter might be, for example:

- The patient themselves, in the context of a consultation;
- A measurement on the patient;
- A trained physician or nurse examining the patient;
- The result of a test performed on the patient.

Do note the distinction between a message from the patient's condition that is some deviation from what is normal, and the signal that may be transmitted by the patient themselves or by a physician or nurse examining the patient. We cannot guarantee that the signal is an accurate representation of the original message (or that the transmitter has not missed a message, or even invented a message).

The channel might be, for example:

- A verbal utterance if transmitter and receiver are in the same room;
- A telephone line;
- An internet link;
- A written communication.

The task of the receiver is to transcribe the signal into an electronic or written record. Finally, the physician is the end point for a sequence of messages that will progressively inform a medical diagnosis.

There are two key points to keep in mind:

1. Any stage in the communication system may lead to loss or distortion of the information in a message as it is transmitted from its source to the physician;
2. Each message will contain a certain amount of *information* that will inform candidate diagnoses.

If we can maximise the amount of information in each message, then we should be able to minimise the number of messages needed in order to reach a confirmed diagnosis.

Performing such a minimisation, of course, presupposes that we do have available some measure of information content. This is the topic of the next section.

2. Information content and entropy

Before introducing a measure of information content, let us first explore a simple example to motivate the precise choice of measure. This is a necessarily brief introduction. A next step for the interested reader may be to read a more extended tutorial such as that provided in [11].

Consider an array of N binary switches, where N could be any integer greater than zero. For each switch, we have two possible states. One could think of these as "on" and "off". Correspondingly, we have two possible messages: one indicating the switch is in state "on" and one indicating the switch is in state "off".

With just one switch, we can store 1 "bit" of information: we just need to receive one message in order to determine the state of that switch. We can store 2 bits of information with an array of two switches, and we will require 2 messages each of 1 bit or one message of 2 bits to determine the state of that array.

Note that our measure of information content is additive; that is, the information content of a single message from an array of 2 binary switches is simply the sum of the information content of single messages sent from each of those switches separately.

In general, with an array of N switches, we will need a message that is N bits long in order to determine the internal state of that array of switches. Now, let us also look at the total number of states of an array of binary switches.

For two switches, we have four possible states: {on, on}; {on, off}; {off, on}; {off, off}. Correspondingly, we will have four possible messages that will tell us the state of the array in a single message.

In the general case of N switches in an array, we have 2^N possible states and 2^N possible messages.

Now, if we assume that each switch acts independently, and that each of these possible messages is equally likely, then any one message has a probability $p = 1/(2^N)$ of occurrence.

Consider an outcome in which a message m of length N is received. The above discussion motivates a requirement for a measure of information content that is additive. In addition, we have seen that the number of states in a system tends to increase exponentially. This suggests the use of a logarithmic function such as that of equation 1:

Eq 1. $$h(x) = -log_2 p(x)$$

Substituting our message m with probability of occurrence $p(m) = 1/(2^N)$ into Equation 1, we get:

$$h(m) = -log_2\left({}^{1}/_{2^N}\right) = log_2(2^N) = N$$

Equation 1 is thus returning us our informally proposed measure of information content; it is in fact the definition of the *Shannon information content of an outcome*.

We now need to generalise this. When a patient presents, that patient's state is not known with certainty. Thus, the possible messages that may be received form an ensemble M, with each message $m \in M$ having a probability of occurrence, $p(m)$. We use the word ensemble here in a statistical sense. Writing this out more formally, M is a triple (m, A_M, P_M) where m is a “random variable” that can take on one of a number of possible values from an alphabet (a set of legal characters) $A_M = \{m_1, m_2, \ldots, m_K\}$ with respective probabilities $P_M = \{p_1, p_2, \ldots, p_K\}$. That is to say, the probability that $m = m_k$ for some $1 \leq k \leq I$ is p_k. We also require that $p_k \geq 0$ for all k, and $\sum_{k=1}^{I} p_k = 1$.

A measure $H(M)$ on the ensemble M can then be defined which is the average Shannon information content of an outcome:

Eq 2. $$H(M) \equiv -\sum_{k=1}^{K} p_k log_2 p_k$$

Strictly, this is simply providing us with the expected value of the information content in a message m that has been received from the ensemble M. However, the form of equation 2 is identical (apart from a constant) to the definition of entropy in the statistical mechanics model of thermodynamics:

$$S = -k_B \sum_i p_i \ln(p_i)$$

Here p_i represents the probability of a certain microstate of the thermodynamic system under consideration, and the sum is over all possible microstates. The natural logarithm is used in thermodynamics, but essentially the different base of the logarithm together with the use of Boltzmann’s constant k_B simply provides a scaling between S and H.

By analogy with the form of this version of Boltzmann’s equation, and the fact that the ensemble M in some sense represents the possible states of the system (a person in our case) under observation, $H(M)$ is referred to as the (Shannon) entropy of that ensemble. As with the Shannon information content, it also has the unit of bits (when using logarithm to the base 2).

Let us look at a couple of general-purpose examples to gain a little more intuition about how Equation 2 might be used before moving back to a diagnostic setting.

Consider an ensemble M in which an outcome is simply a character drawn at random from an English document. That is, the random variable m will be instantiated by selecting at random a character from an English document where A_M = {a, b, c, d, e, …, x, y, z, _}. We will not distinguish upper- and lower-case letters, but we do include the use of a space character, “_”. P_M = {.0575, .0128, .0263, .0285, .0913, …, .0007, .1928}[2] are the respective p_is for $1 \leq i \leq 27$.

Using the figures provided, it can be calculated that the outcome m = “z” has Shannon information content 10.4 bits, while the outcome m = “e” has information content of 3.5 bits. Overall, our English language document has an entropy of 4.1 bits. The full table of probabilities and corresponding measures of information content can be found in [7].

Let us examine this a little more. Providing a clear semantics to Shannon entropy is still a matter of debate (see, for example, p. 65 of [8]). Although it has the same form of thermodynamic entropy, it does not for example have the same units, as we have discussed; equation 2 has units of bits, whilst Boltzmann’s entropy has units of Joules

[2] These values were estimated by the late David Mackay for use in his *Information Theory, Inference and Learning Algorithms* text book, Cambridge, 2003. His choice of text from which to estimate the probabilities, *The Frequently Asked Questions Manual for Linux*, of course means that these probabilities are conditional on the assumption that this text is representative of the distribution of letters in an English language document.

per Kelvin. However, although it would be wrong to say that Shannon entropy is "the same thing" as "entropy", it would be equally wrong to say they are unrelated: the two equations only differ by a constant (which defines the scale of measurement), and one can begin to reconcile the two if one relates the probabilities of the microstates of the system under consideration with the probabilities of the symbols generated by that system. Indeed, Jaynes argued in depth that the information theoretic view of entropy was a generalisation of thermodynamic entropy [3][4]. We implicitly advocate the same position in the context of medical diagnosis.

Going back to our document example. If we take a new document, pick a character at random and that character turns out to be a "z", a character with one of the lowest probabilities of occurrence in a typical English document, then that is providing us with more information about it (relative to a "normal" document) than if we had received an "e".

Two general properties are also worth noting. Firstly, if only one outcome in an ensemble M has a non-zero probability of occurring (in which case, its probability must be 1), then:

Property 1: $H(M) = 0$

(By convention, if $p(m_k) = 0$, then $0 \times log_2 0 \equiv 0$).

At the other end of the scale, the $H(M)$ is maximized if all of the outcomes are equally likely. An expression for the value for this is quite easy to derive. Let our ensemble M_e have K possible outcomes. Then we must have for all k, $p(m_k) = 1/K$. Substituting this into Equation 2, we get:

$$H(M_e) = -\sum_{k=1}^{K} \frac{1}{K} log_2 \frac{1}{K} = \frac{1}{K} log_2(K) \sum_{k=1}^{K} 1 = log_2(K)$$

(Noting that $log(1/K) = -log(K)$ and that $\log(K)$ is a constant and so can be factored to the outside of the summation). So,

Property 2: $H(M_e) = log_2(K)$ if all K outcomes are equally likely

In the case of our English document example, if the characters were uniformly distributed, then we would have $H(M_{uniform}) = \log_2(27) = 4.76$ bits. This is slightly higher than that for our representative English language document (4.1 bits).

Returning to the application of this to medical diagnosis, we can interpret these two situations as follows:

- $H(\ ) = 0$ if only one message/positive test result is possible. That is, a specific diagnosis has been confirmed.
- H is at its maximum when all messages are equally possible. That is, we are at a state of complete ignorance about the patient's internal state.

From this we can see that the challenge of diagnosis is to reduce the entropy to as close to zero as possible, and to select tests so that the result of each test (what we are calling "messages" here) maximises the reduction of entropy.

Two points should be emphasised here before we move on:

1. We are equating the probability of occurrence of messages with the probability of microstates of the patient under examination, to justify the usage of the term "entropy";

2. We are ignoring uncertainties in the veracity of a message that might be introduced by the communication pathway of Figure 2.

Point 2 is critically important when taking this into a real clinical setting. However, for simplicity of exposition we will continue to ignore this issue until the concluding section.

3. Relative entropy and diagnostic tests

Let us phrase the diagnostic strategy a little more formally. A patient's specific internal state has an ensemble of messages $M = (m, A_M, P_M)$ associated with it. A message will normally be triggered by a specific "interrogation" being performed on the patient. An interrogation may be, for example: a question asked of the patient; a test performed on the patient; an inspection performed by a nurse.

Prior to an interrogation the alphabet A_M of messages will have a probability distribution Q_M over it. Receipt of a message m_k (a positive test result, for example) will result in a posterior probability P_M over the alphabet of messages.

To measure the change in entropy, we use the *relative entropy*, or *Kullback-Leibler divergence*, between the two probability distributions [7]:

Eq 3. $$D_{KL}(P_M||Q_M) = \sum_{k=1}^{K} p_k log_2 \frac{p_k}{q_k}$$

It is worth noting two properties of relative entropy. Firstly, it satisfies what is known as Gibbs' inequality, with equality if and only if $P_M = Q_M$:

Eq 4. $$D_{KL}(P_M||Q_M) \geq 0$$

Secondly, in general it is not symmetric under interchange of the two probability distributions. That is, $D_{KL}(P_M||Q_M) \neq D_{KL}(Q_M||P_M)$. Consequently, relative entropy/Kullback-Leibler divergence does not formally qualify as a measure (hence the use of the term "divergence").

Expressed in terms of Bayesian inference, $D_{KL}(P||Q)$ is a measure of the information gained when a physician's beliefs are revised from a prior Q to a posterior P following some investigation.

We will use a hypothetical example adapted from [2] to illustrate the approach so far. We hypothesise a population of patients with arthritis, framed with a prior probability distribution over four possible syndromes. We have two diagnostic tests that are available to us, t_1 and t_2. Table 1 provides the pre-test probabilities and the respective post-test probabilities following a positive outcome from each of the two tests. Which of the two tests provides the greater information gain?

Table 1. Hypothetical Example (adapted from [2]).

Candidate Diagnosis	Pre-test Probability (t_0)	Post-test Probability (t_1)	Post-test Probability (t_2)
Gout	0.25	0.5	0.4
Osteoarthritis	0.5	0	0.1
Pseudogout	0.125	0.5	0.4
Other possibilities	0.125	0	0.1

Using Equation 3, it is straightforward to calculate that the information gain from t_1 is 1.5 bits, whereas the information gain had we chosen to perform t_2 would have been 0.68 bits (to 2 d.p.). Note again, that we make the assumption that the probabilities are continuous and so $0.log_2(0) = 0$. So, in the first case we have:

$D_{KL}(t_1||t_0) = 0.5 \times log_2(0.5/0.25) + 0.5 \times log_2(0.5/0.125) = 0.5 \times 1 + 0.5 \times 2 = 1.5$

In the second case we have:

$$D_{KL}(t_2||t_0) = 0.4 \times log_2(0.4/0.25) + 0.1 \times log_2(0.1/0.5) + 0.4 \times log_2(0.4/0.125) + 0.1 \times log_2(0.1/0.125)$$
$$= 0.4 \times 0.6781 + 0.1 \times (-2.322) + 0.4 \times 1.678 + 0.1 \times (-0.322)$$
$$= 0.68 \text{ (to 2 d.p.)}$$

The question naturally arises: why use relative entropy and not merely the difference of the pre-test and post-test entropies as measured using Equation 2. The latter was indeed proposed in early discussions on the use of entropy in medical decision making. However, Asch, Patton and Hershey concluded that it "fails to capture reasonable intuitions about the quantity of information provided by diagnostic tests" [1]. This point was reiterated in [2], which shows that relative entropy captures those intuitions more effectively. Kullback and Leibler [5], of course, provide a more formal justification of what we are calling relative entropy, as a sufficient statistic for discriminating between two probability distributions.

Let us now take a look at how these concepts from information theory might act as aids in medical decision making.

4. Shannon entropy and binary outcomes

Many laboratory tests are designed to assess the presence or absence of a disease state; a binary outcome. We can take a coin flip as a reference point, with the outcomes being heads or tails. Now, consider a collection of coins that are biased to some extent. That is, each coin will have a probability p that the outcome is a heads, with p varying over the collection between 0 and 1.

For a given coin C, from Equation 2 noting that the probability of a tails will then be $(1 - p)$, entropy is:

Eq 5. $$H(C) = -p \times log_2(p) - (1-p)log_2(1-p)$$

We can see that the entropy varies between 0 and 1, with a maximum at 1 when $p = 0.5$ (see figure 3).

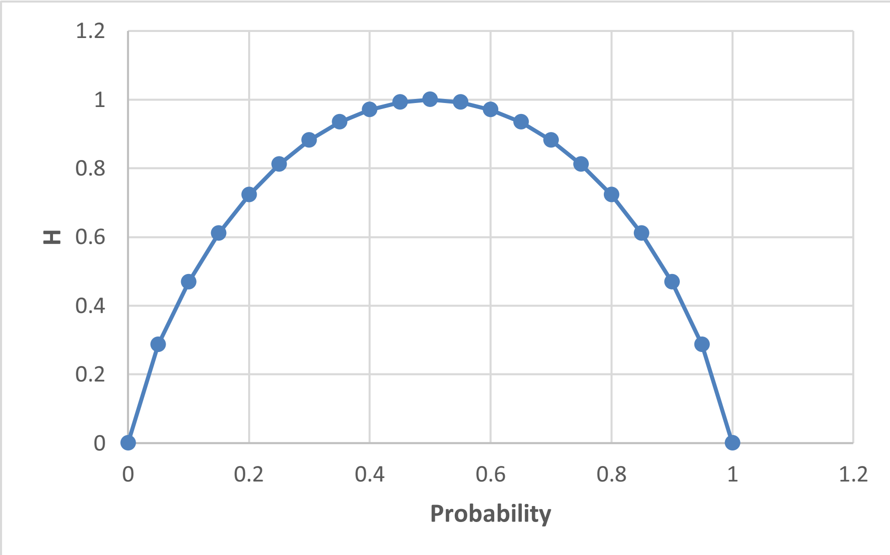


Figure 3. Variation of Entropy vs probability for a biased coin.

Think of the coin flip as a test on the internal state of a patient. A "heads" says the patient may have the disease, a "tails" says the disease is not present. If the coin is unbiased then as a test it is not helping us; all internal states are equally possible. We need a test where the entropy is close to 0 or 1 in order for us to be able to gain anything informative about the internal state of the patient.

Vollmer [13] explored the use of entropy to analyse the information content of a number of laboratory tests. He demonstrated how the concepts from information theory can be used as an aid to evaluating and understanding laboratory test results. We will use just one example to illustrate the point by using Figure 3 as a reference point.

Stadelmann et al [10] reported that the probability of 10-year mortality for malignant melanoma could be estimated from tumour thickness t using the following formula:

$$p = 1 - .966 \times e^{-(0.2016t)}$$

In Figure 4, we plot H vs tumour thickness t, using Equation 5.

We can see that over quite a wide range of values, with median $\approx 3.5mm$, tumour thickness provides limited information about the outcome.

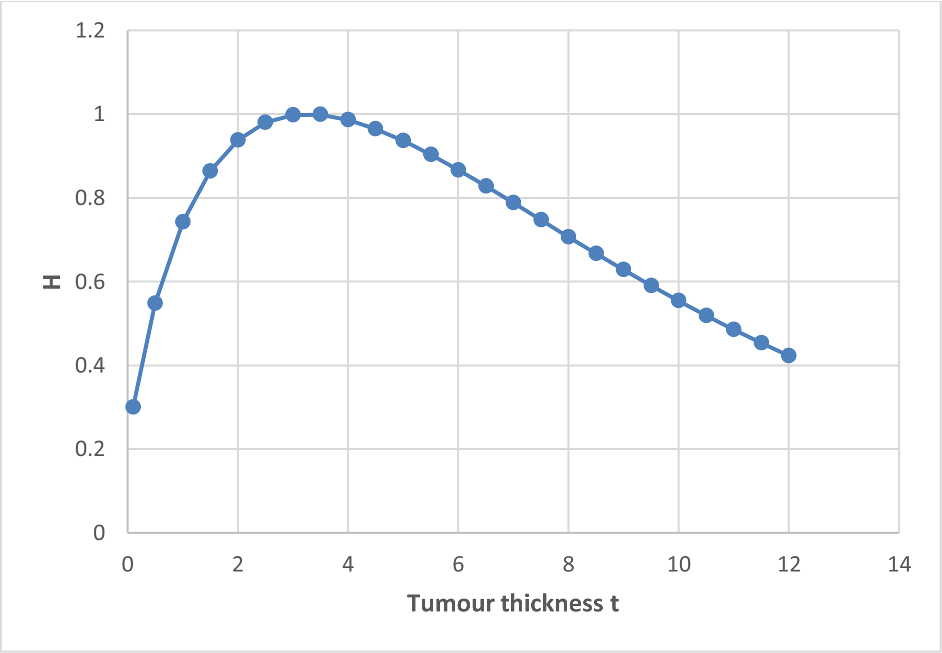


Figure 4. Entropy as a function of tumour thickness *t* (in mm) as a test for 10-year fatality from malignant melanoma.

5. Using information theory to prioritise laboratory tests

We are beginning to build up a number of approaches to using techniques for assisting with the choice of diagnostic tests in a clinical setting. First, we used relative entropy to discriminate between two candidate tests. Then we looked at the variation of entropy with the outcome of some diagnostic test (for a binary outcome) to identify the range of outcomes for which the test is informative. It should be emphasised that these are not the only tools available, but published work to date has argued for their value as additional tools that may aid in decision making.

Lee and Maslove [6] made the case that an information theoretic approach could have particular value in the case of identifying redundancy in tests in an intensive care unit (ICU). Parsimony in the case of blood tests is particularly important in the ICU context as repeated bloodwork can:

Cause anaemia and increase the need for blood transfusions;
Cause patient discomfort;
Disrupt sleep;
Lead to delirium.

The challenge, then, is to identify which blood tests are the most informative at a system wide level. A key issue here is that there may be some level of *redundancy* between laboratory tests; that is, especially over time, some tests may add little information over previously conducted tests. In cases where we can identify that there is a high degree of mutual information between tests (either through the same test being repeated too frequently, of for two different tests having too strong a dependency on common information), then we have an objective basis for reducing the number of tests performed.

Two additional concepts were used in Lee and Maslove [6]. The first of these was the *conditional entropy* of X given Y. This measures the average uncertainty about a random variable x that remains when y is known (x and y being the respective random variables for the ensembles X and Y). It is defined as:

Eq 6. $$H(X|Y) = \sum_{xy \in A_X A_Y} p(x,y) log \frac{1}{p(x|y)}$$

Referring back to the definition of an ensemble, A_X is the alphabet of the ensemble X; that is, the set of legal values of the random variable x. Similarly, A_Y is the alphabet of the ensemble Y.

A related concept is the *mutual information* between X and Y. This measures the amount of information that x conveys about y, and is defined as:

Eq 7. $$I(X;Y) \equiv H(X) - H(X|Y)$$

Note that we have followed the definitions as given in Mackay [7].

Lee and Maslove extracted laboratory test results from MIMIC II, a fully anonymised public database. They analysed a total of 29,149 ICU admissions, investigating the following laboratory tests: haematocrit; platelet count; white blood cell count (WBC); glucose; HCO_3; potassium, sodium; chloride; BUN (Blood Urea Nitrogen); creatinine; and, lactate. Overall, their findings strongly supported the view that a significant amount of the bloodwork performed in ICUs is unnecessary. This had previously been discussed in [14], but Lee and Maslove were able to quantify the level of redundant information content. As a specific example, they found a high level of redundancy in information between the tests for BUN and creatinine; suggesting that if one is known, the other can be inferred with reasonable confidence. Furthermore, their analysis indicated that given the choice, it would be better to prefer BUN over creatinine.

Of course, clinical judgement will always be needed but this information theoretic approach does provide an objective foundation to an informed choice.

6. Discussion

We have shown in this chapter that information theory can have value in informing medical decision making. We have drawn on a number of studies in order to illustrate this. However, there is one area where we do beg to differ with most of those studies. Many of them bring in additional terminology to try and provide an intuitive semantics to some of the concepts in information theory; notions of "surprise", "closeness to certainty", perhaps a tendency to try and equate entropy to uncertainty. One can understand this. Within classical thermodynamics, entropy is perhaps one of the hardest concepts to gain a feeling for. However, we have been careful to refer only to measures of information and entropy. We have briefly alluded to an equivalence between Shannon entropy and entropy from statistical mechanics through associating the messages that can be potentially received from a patient with the internal microstates of that patient.

The paper by Tribus and McIrvine [12] is a good starting point for a deeper study of the conceptualisation of Shannon entropy, and could perhaps be read after the tutorial [11] that was mentioned earlier and the papers by Jaynes [3][4].

We do believe that the time is ripe to propagate a deeper understanding of information theory through practitioners of health informatics. The potential for significant enhancements in the rigour of medical decision making is waiting to be realised. However, some stronger guidelines do need to be developed for its usage. We have described a number of different strategies and there would be real value in documenting a common foundation that would inform all of these. In addition, we would emphasise the need to explicitly model the "noise" that is inherent in the communication model. We have included this in the communication model at the beginning of the chapter, and it is an important but often neglected factor in the risk of misdiagnosis of a patient.

Teaching questions for reflection

1. Can you think of a clinical setting from your own experience where information theory might have usefully informed your choices?
2. Do any of the examples provided in this chapter have more mainstream statistical methods of achieving the same result?
3. What do you feel are the barriers to the adoption of information theoretic approaches in the wider community?

References

[1] D.A. Asch, J.P. Patton and J.C. Hershey, Prognostic information versus accuracy: once more with meaning, *Medical Decision Making* **11** (1991), 45-47.
[2] W.A.Benish, Relative Entropy as a Measure of Diagnostic Information, *Medical Decision Making* **19** (1999), 202-206.
[3] E.T. Jaynes, Information Theory and Statistical Mechanics, *Physical Review* **106** (1957), 620-630.
[4] E.T. Jaynes, Information Theory and Statistical Mechanics, *Physical Review* **108** (1957), 171-190.
[5] S. Kullback and R.A. Leibler, On information and sufficiency, *Ann Math Stat* **11** (1951), 79-86.
[6] J. Lee and D.M. Maslove, Using information theory to identify redundancy in common laboratory tests in the intensive care unit, *BMC Medical Informatics and Decision Making* **15:59** (2015), 1-8.
[7] D.J.C. Mackay, *Information Theory, Inference and Learning Algorithms*, CUP, Cambridge UK, 2003.
[8] L. Pismen, *The Swings of Science – From Complexity to Simplicity and Back*, Springer Nature, Switzerland, 2018.
[9] C.E. Shannon, A Mathematical Theory of Communication, *The Bell System Technical Journal* **27** (1948), 379-423.
[10] W.K. Stadelmann, D.P. Rapaport, S-J. Soong, Prognostic factors that influence melanoma outcome. In: Balch CM, Houghton AN, Sober AJ, et al, eds. *Cutaneous Melanoma.* 3rd ed. St Louis, MO: Quality Medical Publishing; 1998:11-35.
[11] J.V. Stone, *Information Theory – A Tutorial Introduction*, Sebtel Press, 2015.
[12] M. Tribus and E.C. McIrvine, Energy and Information, *Scientific American* **225** (1971), 179-190.
[13] R.T. Vollmer, Entropy and Information Content of Laboratory Test Results, *Am J Clin Pathol* **127** (2007) 60-65.
[14] C. van Walraven and C.D. Naylor, Do we know what inappropriate laboratory utilization is? A systematic review of laboratory clinical audits, *JAMA* **280** (1998), 550-558.

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

doi:10.3233/SHTI190109

Assessing Technology Success and Failure Using Information Value Chain Theory

Enrico Coiera[a,1]
[a] *Macquarie University, Sydney Australia*

Abstract. Information value chain theory provides a straightforward approach to information system evaluation and design. It first separates the different benefits and costs that might be associated with the use of a given information technology at different stages along a value chain stretching from user interaction to real world outcome. Next, using classical decision theoretic measures such as probabilities and utilities, the resulting value chain can be used to create a profile for a particular technology or technology bundle. Value chain analysis helps focus on the reasons for system implementation success or failure. It also assists in making comparative assessments amongst different solutions, to understand which might be best suited for different clinical contexts.

Keywords. Evaluation, Value of information, Utility, Information technology

Learning objectives

After reading this chapter the reader will:

1. Describe the typical steps in an information value chain for information technology use in healthcare.
2. Understand and demonstrate proficiency with value of information calculations.
3. Appreciate that different technologies will generate different value chain profiles.
4. Use value profiles to explain where in the chain a specific technology is most and least effective.
5. Use value chains to diagnose problems in the implementation of a specific technology bundle.

1. Introduction to Information Value Chain Theory

The closely linked processes of design and evaluation are crucial to ensuring informatics interventions work. There are two basic goals when evaluating any information or communication system. Firstly the evaluation must help determine if a system is fit for purpose, or its *efficacy* (i.e. does it do what it is meant to do?). Secondly we may want to decide if a system is the best choice amongst alternatives when used in to solve a problem in the real world (its relative *effectiveness*) (i.e. which solution should we pick?).

[1] Corresponding Author: Enrico Coiera, E-mail: enrico.coiera@mq.edu.au

Neither of these goals is easy to achieve in health informatics, for a number of reasons. Firstly there are many ways to measure success and not every success measure tells us the same thing. Should we pay more attention to surveys of user satisfaction, rates of adoption, or should we only focus on improvements in clinical outcomes? Further, given the complex organizational space within which any informatics intervention must co-exist, it is not surprising the same system, when tested in different clinical settings, usually achieves different outcomes[1]. The effects of the way a technology is implemented, the specific context in which it is used, and how it is used, all shape the outcomes of system use. With so many confounding factors, getting design right or demonstrating system success is non-trivial.

Information value chain theory provides a straightforward approach to both system design and evaluation[1 2]. With its foundations in classical decision science, it provides a mechanism to tease apart the different benefits that might be associated with the use of a given information technology, and also helps identify why expected benefits may not be detectable. Using classical decision theoretic measures such as utilities, and by sequencing the different types of information system functions and their associated outcome measures, value chain analysis helps focus on the reasons for system implementation success or failure. It also assists in making comparative assessments amongst different solutions, to understand which might be best suited for different clinical contexts.

1.1. A value chain extends from system use to health outcome

To undertake a value chain analysis, we begin by describing a value chain that connects use of an information system to final outcomes (Figure 1).

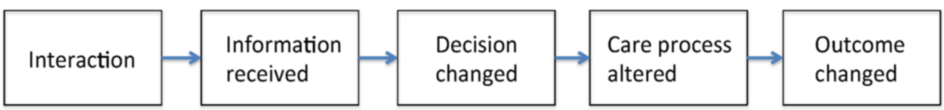


Figure 1. *The information value chain starts with a user interacting with an information system, but must go through many steps before changing clinical outcomes* (from Coiera, 2015).

The chain typically begins with a user interacting with a system (interaction). Some but not all of these interactions potentially providing information that is then received by a user (information received). Some of this information may lead to a decision being changed (decision changed), sometimes leading to a change in the process of care (care process altered). Finally, some of these process changes may then impact the outcome for a patient (outcome changed).

For example, a clinician may interact with an electronic health record, and examine a patient's laboratory test results. Amongst all the information in the available results, one specific test provides important new information. All the other information received adds nothing new and has no impact on the clinician's thinking. Based on that specific new information, the clinician ceases a medication and prescribes a new one (a change in care process). In some circumstances this change is beneficial or harmful to the patient – both leading to a change in outcome for the patient. It is also quite possible that this clinical process change, one amongst many that happen daily, leads to no change in the patient's outcome.

Evaluation can take place at each of these steps in the value chain, but it is not correct to assume that a good result at one step necessarily translates into a good result at the next. Nor should we even desire that improvements at one step in the chain flow downstream. For example, a new telecommunication system between a primary care physician and patients at home may allow a doctor to talk to their patients without the need for a physical visit. We might be able to demonstrate high system utilization and user satisfaction with this telehealth system, but also be surprised to find that there is no significant change to the survival or quality of life for patients. Why might this be so, and should we consider such a system a 'failure'?

There are many reasons why benefits at one step in the value chain do not manifest in later steps. Sometimes a technology intervention behaves as a *substitute* for an existing service process, but does not to improve it. So, when the quality of normal care is already of a high standard, any telehealth substitute for face-to-face interaction should aim to be non-inferior i.e. be no worse. All that we are doing is replacing face-to-face interactions with online ones. If the goal of this telehealth system were only to reduce the need for a patient to travel to the office, then demonstrating a cost-effective reduction in such visits (once we add in the costs of the telehealth system) would be considered a success. There should be no expectation that benefits at the initial interaction stage of the chain translate to clinical outcomes. We should just be mindful to not see any deterioration in outcomes.

1.2. Different evaluation measures may be used at different steps in the value chain

Which stages of the value chain are formally evaluated will depend on the type of system in question, and the purpose of the evaluation. Unsurprisingly, the processes that are studied, and their related measures, can vary with both the step in the chain, and the type of system being developed (Table 1). For example, we might evaluate the quality of interaction with an information retrieval search engine using metrics for the ease with which a query can be formulated to retrieve relevant information, whilst measuring the quality of a telehealth interaction would perhaps focus on the quality of the video call, the rate of technical disruptions to the call, or a user's perceived satisfaction with the call.

1.3. The value of information can be quantified

Value chain analysis makes clear that creating and accessing information alone does not always lead to a change in process or clinical outcome. We know from Shannon's Information Theory that not every additional piece of data is as informative as another [3]. The amount of Shannon information is a measure of how "surprising" new data are compared to our expectation. If data do not tell us anything new, they bring little or no new information. Another way of thinking about this is to ask how many times information must be read before there is a measureable impact on clinical outcomes. Metrics such as the *number needed to read* [4] and *the number needed to benefit from information* [5] are related attempts to correlate access to information such as clinical guidelines with their impact on process or outcome.

Table 1. *Examples of measures that can be used to evaluate systems at different stages of the interaction value chain for information retrieval systems which search for documents, and telehealth systems which support the communication of patient information (n = number of).*

	Interaction	**Information**	**Decision**	**Care process**	**Outcome**
Information retrieval system	*n* queries made, *n* query reformulations	*n* documents retrieved, precision and recall, document relevance	*n* correct or incorrect decisions, decision velocity	*n* and type of tests ordered, medications prescribed, cost of care	Morbidity and mortality, Quality Adjusted Life Year (QALY)
Telehealth system	*n* conversations, call quality and time, user satisfaction	Quality and quantity of patient level data shared	*n* additional correct or incorrect decisions	Health service utilization rates, travel costs	Blood pressure, HbA1c, blood glucose etc., Morbidity and mortality, QALY

Decision theory provides us with a powerful and theoretically robust way of estimating the *value* we place on receiving new information. For example, if a new diagnostic test result changes a patient's treatment and saves their life, then instinctively the value of that information is high. If a diagnostic test allows a patient to avoid a risky treatment and to follow a less risky but equally beneficial option, then the information's value is based on those avoided risks. If a new diagnostic test result only confirms what is already most likely, and it triggers no change to treatment, then it might have a relatively low value.

This *Value Of Information* (VOI) can be defined as the value we place on receiving particular data prior to making a decision [6]. We could calculate such a value in financial terms such as money saved or earned, or as patient expressed preferences. In other words, VOI is the *difference* between the value of persisting with the present state of affairs and the value to us of being able to embark on a new decision, influenced by new information. VOI is zero whenever obtaining new data does not change decisions or outcomes.

VOI also has a decision-theoretic interpretation. Imagine for example that a patient undertakes a test, and will be given different treatments depending on the blood test result. Each of these two treatments will result in a different outcome for the patient. How do we determine the value of each outcome to the patient? A preference for one outcome over another can be represented with a quantitative value called a *utility*. A utility is a number between zero and one and the outcome with the highest utility is the preferred one.

A utility value is thus a model of an individual's preference for an outcome, expressed in numerical form, and can be derived by a number of different means. Common methods to estimate utilities include *rating scales, standard gambles* and estimating quality-adjusted life expectancy e.g. using a *time trade-off* [7] [8].

Next we need to consider that each of the two potential treatment outcomes is uncertain. A given treatment will not always have the same effect on different patients. So even if one outcome might have higher utility for a patient, we need also to consider how likely that utility will ever be realized. To do that we now calculate the *expected utility* e of making one choice over another, which is simply the product of its probability p and its utility u:

$$e(x) = p(x) \times u(x) \quad (1)$$

Expected utility is thus a measure of the actual benefit that can be expected from an event over multiple trials, given uncertainty about the event occurring.

The expected VOI that helps us choose between two different courses of action can now considered to be the difference in the expected utility of the different decision options [9 10] i.e.:

$$VOI = expected\ utility\ (Option\ 1) - expected\ utility\ (Option\ 2) \quad (2)$$

For example, assume that the probability of a clinician finding a new pharmacogenomic test result when interacting with a patient's electronic record is 0.4 because the clinician usually must check the EHR several times before seeing a result. The utility of this result is high at 0.9, because it allows the clinician to choose between two different drug treatments. The Expected utility of this outcome is 0.36 i.e.:

$$0.36 = 0.4 \times 0.9$$

In comparison, the probability of not finding the test result is 0.6. We might assign a utility to proceed without the test result of 0.1 (because there is a good chance that the drug is ineffective for most patients who do not have the gene). The expected utility of proceeding without a gene test is thus 0.06 i.e.:

$$0.06 = 0.6 \times 0.1$$

We can now calculate the VOI for a clinician accessing a gene test result:

$$VOI = 0.36 – 0.6 = 0.3$$

A key idea here is that for new information to have value, the information must be *actionable* in some way. It is not enough that data provide us with a new diagnosis, that diagnosis must then trigger some new action in the world [11]. The action needs to result for example, in a change in morbidity, mortality, or in some other way increase a patient's quality of life. VOI could be negative if the proposed method to gather new information does not lead to an actionable decision with potential benefits, and gathering the data has costs for the patient such as risks of complications that lead to harm, from pain through to injury and even death.

1.4. The value of events along the information value chain can be quantified

Now that we have a way of calculating the value of information for any step in the information value chain, we can turn to look at the way information value changes down the chain. We first look at the frequency with which events occur at each stage in a chain. For example, over a 24-hour period, the EHR in a hospital may be accessed thousands of times, but decision support systems may be accessed only hundreds of times. One interesting property of the information value chain is that there is typically an asymmetry both in the *number of events* at each step, as well as in the *value* of the events (Figure 2).

Firstly, we note that there is a probability for moving from one step in the chain to another. Thus there is a probability (but not a certainty) that interacting with an information system will yield information, or that the information will lead to a decision change. For example, the number of times a clinician reads a patient record is always going to be greater than the number of times that reading leads to a change in decision. Similarly, not every computer generated alert will result in a change in decision. The number of times a decision is changed is also going to be greater than the number of times any such change leads to a measureable improvement in patient care.

Additionally the value of events early on in the value chain will often be lower than for events later on. For example, optimizing user interaction with medication alerts is likely to be of much lesser value than reducing the number of unsafe medication prescriptions, which in turn is of lesser value than reducing the number of adverse outcomes from medication errors. Similarly, the time saved in optimizing a user interaction with an EHR is likely to be of lesser value than improvements to the way tests are ordered, and these are often of lesser value than patient outcome changes such as improved survival or QALYs based on more appropriate investigation of patients.

This typical increase in value of events as we move down the chain is driven by increased value associated with real world health outcome changes compared to the value of improvements in process alone. It is however quite possible that in some settings that it is the early stages in the chain that are of higher value. For example, if human resources are scarce and expensive, then using information tools to optimize human efficiency and effectiveness might have very great value.

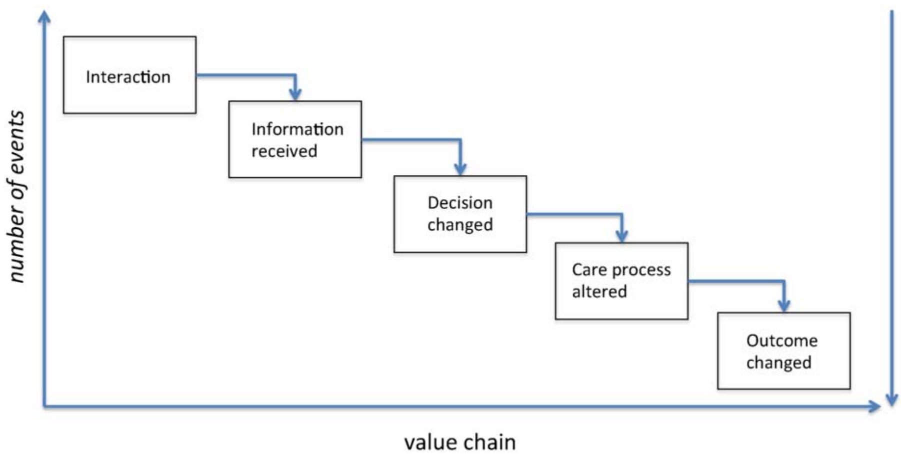


Figure 2: *The number of events is typically higher earlier in the value chain, whilst the value of individual events tends to be higher further down the chain. Combining event frequency (or probability) with event value (or utility) provides the expected utility at each point in the chain* (from Coiera, 2015).

Recall that by combining event frequency (or probability) with event utility, we arrive at an *expected utility*. We can thus calculate the expected utility of using a given system along the different steps in the value chain. The resulting *value profile* of expected utility will not necessarily be constant across the different steps. For example, a telecare system may be designed to maximize expected utility at the interaction stage by reducing face-to-face interactions, but with no expectation of changing clinical outcomes.

A decision support system would be designed specifically to improve decision-making and outcomes, while an EHR is typically designed to improve record keeping, and process improvement goals are reserved for other functions such as CPOE.

We can thus imagine different systems having quite different profiles for their expected utility at different stages of the information value chain. In Figure 3 hypothetical utility profiles are presented for four different classes of informatics intervention. They illustrate that an intervention:

1. May be designed to provide value by improving the quality of interactions in a health service but may provide little additional information compared to current practice (teleconsultation);
2. May optimize the quality of information capture (EHR);
3. May be designed to improve the quality and efficiency of clinical processes (care pathways) or
4. May be intended to intervene in the decision-making process to improve clinical outcomes. Some downstream benefits may even incur an upstream cost (e.g. interacting with some EHRs requires more time than normal practice).

The actual benefits for these intervention classes may be very different, depending on the specific bundle of services offered. For example, it is likely a system that bundles together EHR and decision support will have higher utility than each system alone.

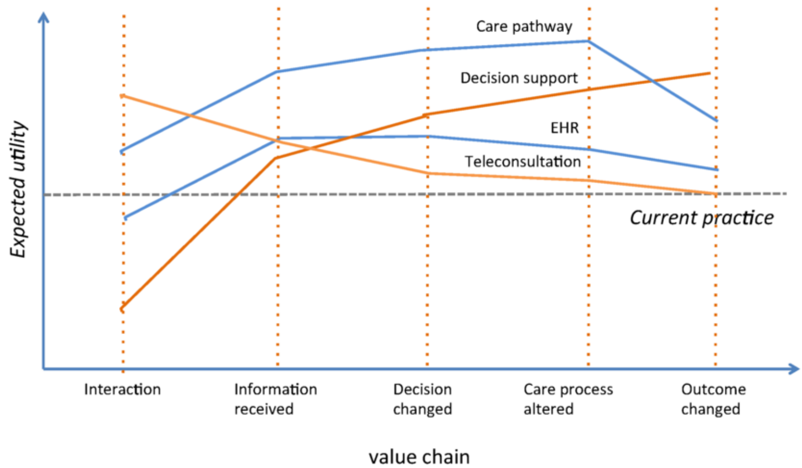


Figure 3: The profile of expected utility for an intervention will vary across the steps of the information value chain, depending on the primary purpose of the system (from Coiera, 2015).

2. Use of information value chain theory in health informatics

Given its relative simplicity, and its foundation in standard decision theoretic concepts such as utility and value of information, value chain analysis has broad application in healthcare. In particular, it can be used to assess the specific benefits of a given technology, or make comparative assessments between competing technologies. Such evaluations might happen post-hoc, for example trying to explain why outcomes for a particular technology implementation did not meet expectations. They can also be used much earlier on, in system design, when the likely impact of different technology bundles is compared and decisions made about system design.

2.1. Case Study 1: The value of using national summary electronic health records

For any formal evaluation of an electronic record system (EHR), whether at a single institution or at nation scale, measurements need to be taken at multiple points along the value chain (Figure 4). The outcome at any stage can only be understood by modeling earlier upstream events. Thus, failure to demonstrate clinical outcome changes following the implementation of an EHR might arise because of problems with events early in the chain e.g. record quality. Alternately, a lack of impact on outcomes may be unrelated to the EHR (for example organizational challenges may prevent important information from the EHR being translated into process changes).[12]

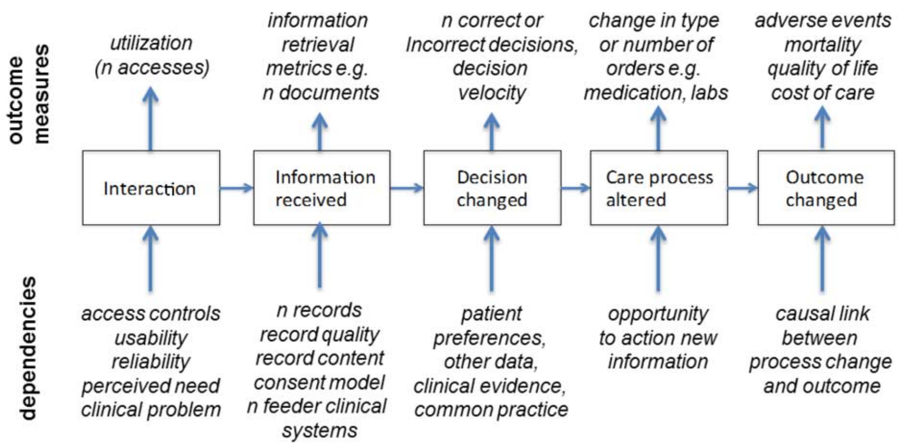


Figure 4*: The information value chain provides a simple causal model connecting EHR use and clinical outcomes. Each step is characterized by different measures, and is dependent on different elements of shared record system design and use (adapted from Bowden and Coiera, 2017).*

For example, imagine that a government has built a national summary health record for every citizen. The system is classed as a success because a large number of citizens have records created for them, and there is a regular stream of record updates every month. What if we however look not at how much data are uploaded into the system, but how often clinicians queried the data? If the system was not often used to support clinical care, perhaps the evaluation might be very different.

Evaluation might reveal that the system was not easy to use by clinicians (who therefore were abandoning it), or that the information within the records was not useful, or even that the systems in place to access the records were not mature compared to the data upload arm of the system. Finally one might look at the downstream impact of system use on the cost and quality of care delivered. What changes to care result from accessing the record? Do these changes translate into better decisions that improve patient outcomes or create service efficiencies? It might prove very difficult for a government to answer these final questions, and very easy to provide data about record or usage numbers. There is however no logical reason to assume that usage of a system translates into changes in end outcomes.

2.2. Case Study 2: Clinical Audit and Feedback

Audit and feedback (A&F) interventions have had mixed success in ensuring patients receive improved care [13 14]. Unlike clinical decision support tools, which provide clinicians with patient-specific advice at the point of care, A&F tools provide data about quality indicators at a population level over a period of time. Reasons for their variable effectiveness are unclear because the mechanisms behind intervention success or failure are poorly understood [15].

Value chain analysis can assist in identifying where potential barriers to effective use of A&F reside[2]. For example, in a situation in which A&F is focussed on improving prescribing, does the type and number of feedback alerts a clinician receives influence the probability that clinicians actually notice them, or subsequently influence their decision making, or which medications are dispensed by pharmacists or finally how many unscheduled hospital admissions are prevented?

In a study by Gude et al., the number of events at each stage of the A&F value chain for medication prescription were measured [16]. System designers were faced with a situation in which A&F was not having any perceptible impact on clinical outcomes, and wanted to understand why this was the case. Analysis of the A&F value chain (Figure 5) reveals a major disconnect between events. Firstly there is a steep reduction between the number of indictors demonstrating poor performance, and the number of indicators flagged for action. An even more dramatic reduction occurs between the problems identified by these indicators, and any action to change clinical process. Of 379 indicators targeted for an action, only 31 were addressed. The study noted “feedback did not lead to teams focusing their quality improvement decisions on low performance areas, and that planned improvement actions were often not completed”.

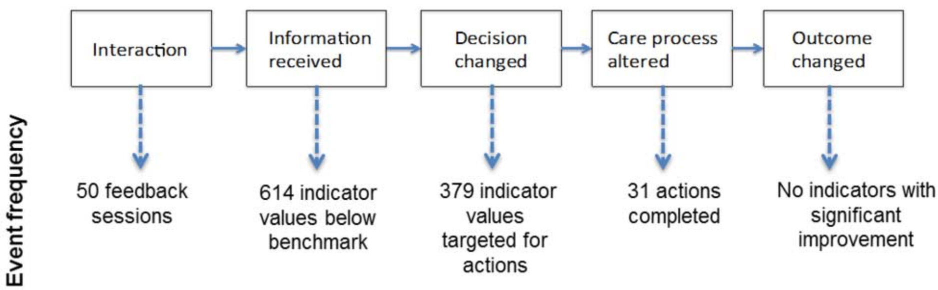


Figure 5. *The information value chain for a computerized Audit and Feedback (A&F) intervention in cardiac rehabilitation. Clinical teams received feedback multiple times on a set of eighteen quality indicators (adapted from Gude et al. 2016 [16]).*

Focusing just on the probabilities of events in the value chain, as shown in Figure 5, can tell us *where* a problem is occurring. The next stage of analysis requires measuring the utility of events at each step, to provide more focused information on *why* events do or do not occur. In this case study, measuring utility can help identify the source of the problem more precisely. Was the lack of outcome change because the alerts about abnormal indicators (information received), were of low perceived value (perhaps

[2] See Chapter 14 “Control Theory to design and evaluate audit and feedback interventions” for an analysis of the same case using Control Theory.

because there were too many of them, or the alerts had low sensitivity or specificity). Was it instead that the cost of changing a clinical process (care process altered) was too high, perhaps because clinical staff were resource constrained, and had little capacity to make the changes needed?

Table 2 provides an example of the calculations that can be made for expected utility, based on measurements of the probability and utility of each step in a value chain. It demonstrates that in this particular scenario, the problem lies in the implementability of decisions to improve practice. There is clear benefit in what the Feedback and Audit tool tells clinicians, and it is also clear that there is benefit in undertaking the recommended changes. There however is no ability to translate this feedback into effective real world actions. The main problem in this example is not with the technology, or the information it generates, but with the socio-technical context in which it is used. Consequently creating a better tool would still not change the outcome. Instead, more resources and leadership might be needed to action the information generated by the analytics tool.

Table 2. *Worked example of a value chain analysis for a computerized Audit and Feedback report. Probabilities are obtained by measuring real world event frequencies, Local utilities are obtained by measuring clinician value assessments at each step in the value chain, using a standardized measurement instrument. The expected utility for any path fragment is calculated from the utility of the node at the end of the path and the probabilities of every node in the path.*

	Step 1: Interaction	**Step 2: Information received**	**Step 3: Decision changed**	**Step 4: Care process altered**	**Step 5: Outcome changed**
Event probability	1.0 (1000/1000)	0.61 (614/1000)	0.62 (379/614)	0.08 (31/379)	0 (0/31)
Utility	0.8	0.9	0.9	0.92	0.95
Local expected utility	0.8 (0.8 x 1.0)	0.55 (0.9 x 0.61)	0.56 (0.9 x 0.62)	0.074 (0.92 x 0.08)	0 (0 x 0.95)
Path expected utility	0.8 (0.8 x 1.0)	0.55 (0.9 x 1.0 x 0.61)	0.34 (0.9 x 1.0 x 0.61 x 0.62)	0.028 (0.92 x 1.0 x 0.61 x 0.62 x 0.08)	0 (0.95 x 1.0 x 0.61 x 0.62 x 0.08 x 0)
Analysis	Utility of a accessing 1000 indicators is high because the there is a high expectation they will contain actionable information. Report length may reduce utility.	Utility of receiving specific information from a report about abnormal indicators is high, but expected utility is lower as probability that any indicator is abnormal is moderate.	Utility of decision to deal with an abnormal indicator is high, given likely benefit. Expected utility is lower as only some indicators are chosen.	A collapse in expected utility at this stage occurs because most decisions in Step 3 do not translate into process changes in Step 4.	The potentially high utility of process changes is entirely negated by the very limited process changes arising from Step 4.

3. How value chain analysis can assist in explaining health IT success or failure

Value chain analysis has a role both in explaining what has already occurred, through retrospective evaluation, as well as in shaping the design of technology and the way it is embedded in the larger socio-technical system. It can be applied in a number of different circumstances, including:

1. *Qualitative retrospective analyses*. The overall evidence for the benefit of a specific technology is often patchy, and the choice of outcome measures for evaluations may not always be ideal. As we saw in Case Study 1, it is easy to pick intermediate process measures which give a false sense of success, just as it is easy to over-emphasize clinical outcomes when the real benefit of a technology is to optimize events earlier in the chain. Value chain analysis can provide a template to consider the real-world costs and benefits of a technology at different points in the chain, identifying gaps in knowledge about performance, as well as guiding the interpretation of success and failure [17].
2. *Quantitative retrospective analyses*. When performance data are available for a specific system, as in Case Study 2, then value chain analysis can reveal specific problems in the design, implementation or use of a system. Event frequency data is ideally recorded automatically as part of system operation, and utility data can be obtained from system users, potentially even retrospectively.
3. *Prospective quantitative studies*: If a value chain can be provided with estimates of expected usage and benefit of an implemented technology, it can be used to provide predictions about overall system utilization and benefit. Such hypotheses can then be tested in prospective trials.
4. *Technology design*: Typically a digital service is built up of a bundle of separate elements. A decision support system bundle will actually require components that access the electronic health record, a user interface, and alerting strategy, and so on [18]. The overall performance of the bundle is thus dependent on the performance of individual components, and the dependencies between components. For example, if the electronic record component is suboptimal, then it does not matter how good the decision support engine might be, as the quality of recommendations will still be poor quality. System designers can estimate the necessary value profile for each element of a bundle, so that together the bundle performs as expected.

4. Discussion

Value chain theory makes very few assumptions about the nature or purpose of technology, and so has broad applicability. The strongest assumption is that the purpose of technology is to improve specific decisions, and that there is a prospect that those decisions have a detectable outcome in the real world. By relying on standard tools such as probabilities and utilities, value chain analysis is strongly grounded upon well-accepted and proven analytic concepts and methods.

One can consider a value chain to be the equivalent of a single path down a decision tree, but with some key differences. Most critically, in a decision tree we only calculate the utility of the final or terminal node. What is interesting about value chains compared to decision trees is that each node in a chain *could be* the terminal node, each with its own intrinsic and different utility in the world. One could stop a chain at reading an electronic record, and calculate the expected utility *to this point only*. Alternatively, one could add decision support to the electronic record, which would change the utility and disutility associated with system use. Since some electronic records have decision support, and some do not, these separate calculations of utilities allow us to make comparisons using the value chain. For a decision tree, we calculate expected utility by multiplying the utility of a terminal node by the probabilities of each step in the path to that node. We calculate a similar path expected utility in a value chain, but can do so for each node in the chain (see Table 2). This path expected utility for a node in a value chain represents the *expected utility of ending the chain at a given node*.

A related question is whether the utility of one node directly determines the value of the subsequent nodes. The answer is that earlier nodes in a chain do influence the utility of later ones, but not in an easily definable way. A value chain is typically an open world. Each node has a separate utility because different populations of patients and users, technologies and external factors all might contribute to each node's utility. So whilst each earlier stage does shape downstream utility, we do not know the specific mathematical function that describes how it contributes, and there is no easy way to infer one directly from the other. For this reason we re-measure utilities at every node.

Although value chain theory is essentially quantitative – it asks us to calculate the value of information at different steps – it is important to remember that in many cases we will be making qualitative comparisons between different stages in the chain. This means that in some cases where great precision in value calculation is difficult, approximating the value of information still allows meaningful qualitative comparisons to be made – usually where there is substantial difference in the VOI at different stages in the chain. As with any theory that relies on quantitative measurements, it is important to ensure that data used in any analysis actually measures what it is meant to. Standard epidemiological challenges such as dealing with confounding factors and noise, as well as temporal variations such as seasonality in disease and service patterns, all need to be addressed.

It is important to recognize that value chain theory does not attempt to provide detailed mechanistic explanations for the impact of information technology beyond the causality implied in the structure of the chain itself. From this perspective it provides a lens to focus on areas of concern or benefit, and other approaches to analysis that assist in untangling the reasons for a particular outcome are then needed.

Value chain theory can also help answer questions about the need for automation, and thus help decide which tasks should or should not be automated [19]. Recognizing that there will likely be different expected utility profiles for completing a task by machine or by human, we can calculate both profiles and plot the resulting curve to generate a summary profile (Figure 6). Undertaking this type of analytic exercise allows us to identify whether tasks are better automated, left to humans, or performed jointly [2]. Understanding the answer has fundamental implications for the strategy taken and its likelihood of success.

Whilst the generic value chain in Figure 1 is applicable to a broad class of information and communication systems, there appears to be no theoretical restriction to imagining different chains of events, or adapting this chain to meet the needs of a specific setting, technology or purpose. One alternate formulation by Parasuraman et al. uses a simplified four step information processing model to create a similar pipeline [20], in contrast to the model used here, which is instead based on human decision making.

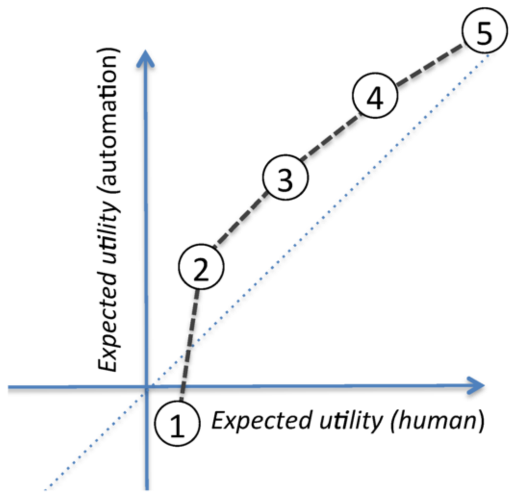


Figure 6: The expected utility (EU) of completing a given task by human or computer can be plotted over a task space. The information value chain (represented as 5 separate tasks) can be plotted into this human-computer task space. The resulting value profile is a function of the given task, the specific technology implementation, the human user, and the context of use. The shape of the plot will likely vary by changing any of these four variables (from Coiera, 2016).

Teaching questions for reflection

1. Describe the typical steps in an information value chain and explain how you would measure the effectiveness at each step for a conversational agent that assists patient's check their symptoms and decide whether to seek professional help.
2. What is the value of information for a new radiological test that has an accuracy of 95%, and which is 20% more accurate in identifying early stage cancer than the current standard test, knowing that undetected cancers will otherwise result in death? Patients report that on a scale of 1 to 10 for discomfort, the new test rates 6, whilst the old test rated 2.
3. Figure 3 shows possible value profiles for several different classes of health information system. Which bundle of two technologies is most likely to improve patient outcomes, using three example profiles?
4. Looking at the value profile for telemedicine, how reasonable is it to expect that widespread use of telemedicine will improve patient outcomes? Contrast the scenario where patients all have full and easy access to face to face consultations with the circumstance where patients are in remote settings.
5. What advice would you give to a hospital proposing to implement a new Audit and Feedback tool to improve the quality of their oncology service, referring to the experience in Case Study 2?

Acknowledgment. This chapter is based in part on material originally published in the Guide to Health Informatics, Third Edition, (2015) published by CRC Press.

References

1. E. Coiera. *Guide to Health Informatics (3rd Edition)*. 3rd ed. London: CRC Press, 2015.
2. E. Coiera. A New Informatics Geography. IMIA Yearbook 2016:251-55
3. C.E. Shannon, W. Weaver. A mathematical theory of communication: American Telephone and Telegraph Company, 1948.
4. B Toth, J Gray, A Brice. The number needed to read—a new measure of journal value. *Health Information & Libraries Journal* 2005;**22**(2):81-82
5. P Pluye, R.M. Grad, J. Johnson-Lafleur, et al. Number Needed to Benefit From Information (NNBI): Proposal From a Mixed Methods Research Study With Practicing Family Physicians. *The Annals of Family Medicine* **11(6)** (2013), 559-67.
6. R.A. Howard. Information value theory. *Systems Science and Cybernetics, IEEE Transactions* **2**(1) (1966), 22-26.
7. M.G.M. Hunink, P.P. Glasziou, J.E. Siegel, et al. *Decision making in health and medicine - integrating evidence and values*. Cambridge, UK: Cambridge University Press, 2001.
8. H. Sox, M. Blatt, M. Higgins, et al. *Medical Decision Making*. Stoneham, MA, USA: Butterworths-Heineman, 1988.
9. A decision analytic method for scoring performance on computer-based patient simulations. *Proceedings of the AMIA Annual Fall Symposium;* 1997. American Medical Informatics Association.
10. E. Fenwick, K. Claxton, M. Sculpher. The value of implementation and the value of information: combined and uneven development. *Medical Decision Making* **28**(1) (2008), 21-32.
11. R.S. Braithwaite, M. Scotch. Using value of information to guide evaluation of decision supports for differential diagnosis: is it time for a new look? *BMC medical informatics and decision making* **13**(1) (2013), 105.
12. T. Bowden, E. Coiera. The role and benefits of accessing primary care patient records during unscheduled care: a systematic review. *BMC medical informatics and decision making* **17**(1) (2017), 138.
13. N. Ivers, G. Jamtvedt, S. Flottorp, et al. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Systematic Reviews* 2012;**6**(6)
14. P.S. Roshanov, S. Misra, H.C. Gerstein, et al. Computerized clinical decision support systems for chronic disease management: a decision-maker-researcher partnership systematic review. *Implementation Science* **6**(1); (2011), 92.
15. N.M. Ivers, A. Sales, H. Colquhoun, et al. No more 'business as usual'with audit and feedback interventions: towards an agenda for a reinvigorated intervention. *Implementation Science* **9** (2014), 14
16. W.T. Gude, E. Coiera, N. Peek. Optimizing Digital Health Informatics Interventions Through Unobtrusive Quantitative Process Evaluations. *Studies in health technology and informatics* **228** (2016), 594-98.
17. M.O. Kim, E. Coiera, F. Magrabi. Problems with health information technology and their effects on care delivery and patient outcomes: a systematic review. *Journal of the American Medical Informatics Association* **24**(2) (2017), 246-50.
18. A.Y.S. Lau, J. Proudfoot, A. Andrews, et al. Which bundles of features in a web-based personally controlled health management are associated with consumer help-seeking behaviors for physical and emotional wellbeing? *Journal of Medical Internet Research* **15**(5) (2013), e79.
19. V. Sintchenko, E.W. Coiera. Which clinical decisions benefit from automation? A task complexity approach. *International journal of medical informatics* **70**(2) 2003, 309-316.
20. R. Parasuraman, T.B. Sheridan, C.D. Wickens. A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics* - Part A: Systems and Humans **30**(3) (2000), 286-97|.

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

doi:10.3233/SHTI190110

Linking Activity Theory with User Centred Design: A Human Computer Interaction Framework for the Design and Evaluation of mHealth Interventions

Alice GOOD[a,1] and Omobolanle OMISADE[a]
[a] *School of Computing, Faculty of Technology, University of Portsmouth, UK*

Abstract. mHealth can offer great potential for the self-management of health conditions and facilitating health services. It is therefore imperative that the design of mHealth systems afford optimum efficacy and effectiveness. Involving end users in collaborative decision making is an essential aspect of increasing acceptance of the treatment intervention. Involving users in the design and evaluation of mHealth systems helps to enable a better understanding of the complexity of user needs and how to incorporate this information effectively into the design process. This chapter discusses how Activity Theory can help to provide a theoretical lens for a User Centred Design framework in the design of mHealth systems. A general overview of Activity Theory and User Centred Design are provided, followed by their application in mHealth. Two use cases are provided that demonstrate how Activity Theory has helped provide a broader contextual analysis to a User Centred iterative approach to system design and evaluation.

Keywords. User Centred Design, Activity Theory, Human Computer Interaction, mHealth

Learning objectives

After reading the chapter, the reader will be able to:

1. Understand the groundings of Human Computer Interaction, User Centred Design and Activity Theory.
2. Demonstrate how Activity Theory can be applied to a User Centred Design approach throughout the technology life cycle design.
3. Understand how to apply User Centred Design principles in the design, implementation and evaluation of mHealth applications.
4. Explain how Activity Theory can provide a contextual analysis of user needs in mHealth applications.

[1] Corresponding Author: Alice Good; E-mail: alice.good@port.ac.uk

1. Introduction to User Centered Design and Activity Theory as a Human Computer Interaction Framework

In recent years, there has been a huge increase in the development of mHealth interventions to improve healthcare delivery and services. The ubiquity of mobile devices allows for the provision of mHealth interventions, which affords several benefits including patient autonomy through self-management, cost saving and an increase in health literacy [2, 3]. To optimise both the effectiveness and usefulness of these mobile interventions, usability and acceptance are paramount. Involving end users in the technology lifecycle process could help to ensure that end users' needs, and expectations are met, as well as increasing the likelihood of acceptance and adoption for optimum clinical impact where relevant [4]. This section explores the origins of User Centred Design (UCD) and Activity Theory (AT), and explains the usefulness of their application in terms of understanding the complexities of users and their interactions in system design.

1.1. An overview of Human Computer Interaction

While there has been a significant development in mHealth self-management interventions, these are sometimes lacking in theoretical underpinnings and adequate assessment of end-user needs [5,6], which then restricts their effectiveness. Human Computer Interaction [HCI] is an interdisciplinary field concerned with the design, implementation and evaluation of interactive systems [7]. Its methodologies and theories are drawn from multiple fields including computer science; sociology; psychology; ergonomics; anthropology and cognitive science, however its roots fundamentally lie in the social sciences, specifically cognitive theory and human factors [8]. The theoretical underpinnings of HCI contribute to the understanding of aspects of design specifically relating to perception; cognition; behaviour and interaction. A selection of some of the theories and models used within HCI include:

Perception (Psychology)

- Hick's Law: states the time it takes to make a decision increases as the number of alternatives increase [9].
- Fitts' Law: predicts that the time required to rapidly move to a target area is a function of the ratio between the distance to the target and the width of the target [10].

Cognition & Behaviour (Sociology)

- Action Theory: Norman's seven stages of action models the way people act when they are interacting in the world to achieve their goals [11].
- Activity Theory: a theoretical framework for analysing human practices as developmental processes with both individual and social levels interlinked at the same time [12].

Interaction (Unique to HCI)

- GOMS: models tasks and user actions; set of Goals, Operators and Methods for achieving goals, and a set of Selection rules for choosing methods for goals used [13].
- KLM: predicts how long it takes a user to complete a task. Based on GOMS, it provides an analysis of steps taken [14].

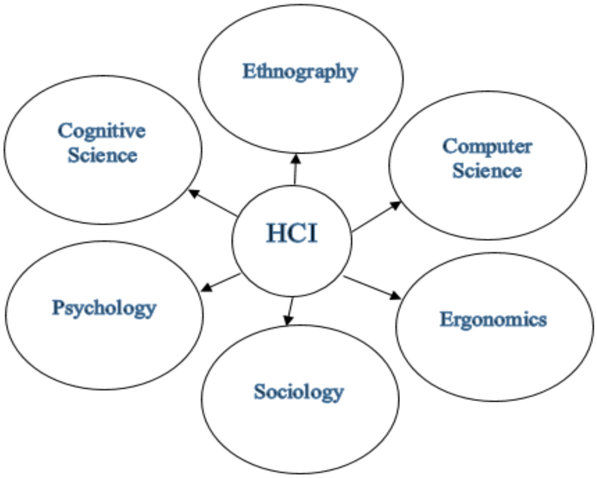


Figure 1: The interdisciplinary field of HCI

Human decisions that relate to health outcomes are multifaceted and are enacted within complex and dynamic contexts. HCI offers a means to enable designers to comprehend how humans use and interact with health systems [15]. Engaging users in the design process early and throughout the design life cycle helps to develop systems that are easy to learn, increase user productivity and satisfaction, increase user acceptance, decrease user errors, and decrease user training time [7]. Several design methodologies for HCI focus on feedback and conversation between users, designers and the technical system. Furthermore, research has identified that healthcare researchers, software developers and practitioners often overlook relevant user characteristics, user tasks, user preferences, and usability issues, resulting in systems that decrease productivity or simply remain unusable [7,16]. The importance of involving target users in the design process for effective interaction with mHealth interventions is, therefore, emphasised [8]. mHealth interventions need to be developed with adequate consideration of the needs of their intended users so that they are efficient, easy to use and perceived as useful [15]. This has increased the interest in applying a UCD approach to mHealth interventions [17,18].

1.2. Understanding User Centred Design as an approach to successful design, implementation and evaluation of interactive computer systems

UCD is an approach that places users at the centre of the design process from the stages of planning and designing the system requirements to evaluating and deployment of the product [18]. UCD refers to how end users influence design through their involvement in the design processes and has been shown to contribute to the acceptance, adoption and success of systems [2]. It can be characterised as a multistage problem-solving process that not only requires designers to analyse and foresee how users are likely to use a product, but also to test the validity of their assumptions with regard to user behaviour in real-world tests with actual users. Figure 2 shows the iterative stages of UCD.

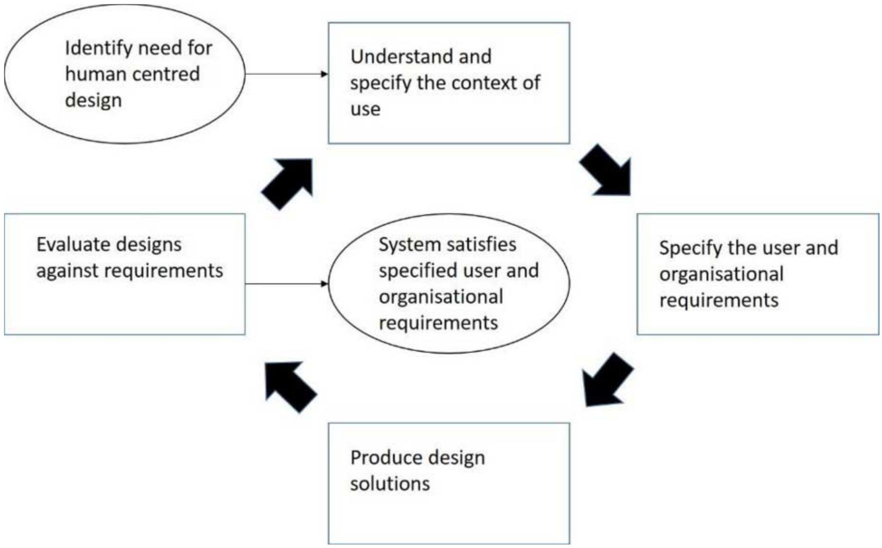


Figure 2: ISO 9241-210 standard for human-centred design processes for interactive systems (Based on [19])

ISO (the International Organisation for Standardisation) 9241-210 provides clarification on the principles of human-centred design and how human-centred methods can be used throughout the system life cycle. These principles are described as follows:

- **Understand and specify the context of use**. This consists of gaining a clear understanding of the users, task analysis, as well as context and environmental analysis. Each stage of this analysis can be dynamic and provides different but necessary components that inform the design of the system. The user analysis consists of examining and learning about the characteristics of the intended users [7]. The task analysis examines the goals of the user, the required functionalities of the system. The environmental analysis not only examines the environment in which the users work but also their social and cultural milieu [14]. It specifies the conditions in which systems are used.
- **Specify the user and/or organisational requirements**. This can be achieved through various ways such as including end users in a design team, as well as consulting with potential end users and relevant stakeholders to assist in requirements usability testing. Participatory design involving end users is an important component of UCD and should be upheld throughout the design and development process of a system [19]
- **Produce design solutions**. Findings from the evaluation inform the design and implementation of the system. This principle emphasises the importance of user centred evaluation to inform the design and to improve it within all stages of the technology life cycle. Prototyping from low fidelity (paper prototypes) through to high fidelity and modelling interaction and tasks can be adopted to design and evaluate the system. Storyboarding, which facilitates the communication of design to potential users, post-experience interviews and satisfaction questionnaires for preliminary design can also be used [19].

- **Evaluate designs against requirements**. In accordance with the UCD ISO process, evaluation of the iterative interface design is evaluated against the requirements. This iterative design is a way of getting end users involved in the process. This includes active user involvement in evaluation and design throughout the entire development process, and the evaluation of use in the context of real user goals and environments [8]. The system is continuously evaluated, using the results to inform the requirements of system redesign where necessary. A range of methods including empirical user testing, heuristic evaluation and cognitive walkthroughs are applied. Gaining end-user feedback is an integral part of this stage [19].

The design addresses the whole user experience, not solely focusing on usability but also promoting a positive user experience during the interaction. User experience can be evaluated through the use of interviews and/or questionnaire which probe end user experiences after using a system [18]. The design and development team includes multidisciplinary skills and perspectives. This can be a combination of various perspectives that can include the stakeholders, potential end users, experts, non-technical. Several techniques used within UCD are shown in Table 1 below [14-16].

Table 1. Overview of the User Centered Design framework.

UCD techniques	Stages in User Centered Design cycle			
	Context of use	**Requirements**	**Design**	**Evaluation**
Diary Study (medical research)	✓			
Activity theory	✓	✓	✓	✓
User Personas	✓			
User scenarios	✓			
Focus groups	✓			
Interviews	✓		✓	✓
Participatory design		✓	✓	✓
Paper prototyping		✓	✓	
Empirical evaluation			✓	
Cognitive walkthrough			✓	
Think-aloud protocol			✓	
Heuristic evaluation			✓	
Surveys	✓		✓	✓
Field studies	✓			✓

1.3. Limitations of User Centred Design in providing contextual user analysis

A usable application is one that not only understands the fundamental needs of the user, but also understands a user's situation, i.e. context and environment and takes appropriate actions to enable the user's tasks. To enable this, the application needs to collect and infer relevant contexts to understand the user's situation. However, this is limited when applying UCD solely to the development of mHealth interventions. Users will invariably have different perceptions, understanding and expectations of a mHealth system and this will affect how they interact with the system. Furthermore, human activity is directly influenced by social, cultural and historical context [12], which adds further complexity. It is therefore important to consider the way people interact with mHealth interventions and their perception of the purpose of using or engaging with the mobile intervention. The necessity of understanding the users and context of use of health systems has already been emphasised. It is then necessary that the social, cultural and psychological aspects of the user in context is captured [21]. One method of achieving this is by using a theoretical tool such as Activity Theory [AT], which can help provide emphasis on user context and interaction. Applying AT to UCD can help bridge the gap in adapting contextual information to the user's situation and needs.

1.4. Introducing Activity Theory as a conceptual framework in User Centred Design.

AT is a theoretical framework for analysing human practices as developmental processes with both individual and social levels interlinked at the same time [12;22]. AT was originally based upon the work of Vygotski and the study of cultural-historical psychology [12]. The AT framework uses 'activity' as the basic unit for studying human practices. Activity or 'what people do' is reflected through actions as people interact with their environment. It can be conceptualised and used in a variety of ways. Because of the way in which it can provide a richer analysis of user needs and context, AT has been adopted as an HCI conceptual framework to help guide and inform the different stages of UCD. Our understanding of AT can be described as follows: the basic unit of analysis is an activity which includes a context. An activity includes eight components with a triangular relationship.

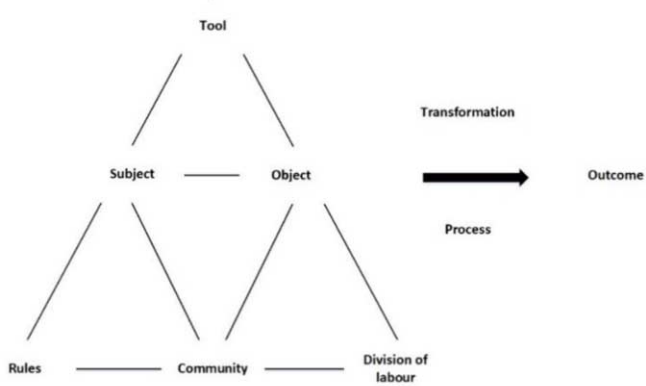


Figure 3: Engeström's Activity system model (Based on [23])

As illustrated in Figure 3, the components of activity include subject, object, tools, rules, community, division of labour, transformation process and outcomes. Engeström modified the original activity framework to include 'rules and division' of labour to understand work activities. An activity is bounded by its subjects (individual or subgroup) and objects (objective). Activities are directed at objects [or tasks] and are motivated by the need to transform the object into an outcome [12]. The relationship between the subject and object of the activity is mediated by a tool or instrument. Tools are the means used in performing an activity; they can enable the subject to transform objects into outcome. This is located within a community, and the community is governed by a set of rules and organised through a division of labour. AT can be applied at different stages in the UCD life cycle. It is particularly effective in capturing the requirements for a computer system design to establish what it is the end-user wants from the proposed computer system [22]. AT has also been used for the design of context-aware applications [23] because it is fundamentally user centred in its approach.

2. Linking Activity Theory with a User Centred Design approach in mHealth interventions

This section provides an overview of how an HCI approach can be used in the design of mHealth interventions, by using AT to provide a conceptual framework when using the UCD framework. Firstly, the importance of a UCD approach in mHealth is highlighted, along with the theoretical framework that AT provides. Two use cases are then provided to illustrate how AT can be linked within a UCD framework and then applied throughout different stages of the reiterative project lifecycle. In the first use case, AT and UCD are applied in the initial three stages: requirements analysis, design and evaluation, whereas the second use case provides an example of usage in the deployment stage and then returns to the requirements stage following analysis.

2.1. Overview of how User Centred Design has been used in mHealth Interventions

UCD begins with a thorough understanding of the needs and requirements of the users. This is critical to the success of mHealth interventions. Understanding user requirements can improve user satisfaction and user experience, increase acceptance and adoption rates and reduce the need for training [24]. The role of user requirements analysis in the development of healthcare interventions is fundamental [7, 17]. Users are generally not brought into the developmental process until after the design brief for a new product has been produced, which can lead to assumptions being made and ultimately lead to the failure of systems [15, 7]. It is recommended that patients are actively involved in the design of healthcare technologies, to help promote a better match to user needs and increase acceptance. UCD provides an approach that enables end users to participate during the life cycle of technology intervention. UCD does this by adhering to core principles that were outlined in section 1.

AT has been used to provide a theoretical framework for collaborative decision making in analysing mHealth systems [25]. The process of engaging patients in decision making can help provide a positive impact upon treatment adherence and health outcome. AT can be applied at different stages in the project life cycle. The two use cases presented below give examples of how activity theory has been applied to a user centred design framework of implementing mHealth interventions.

2.2. Use case 1: The design of a Mobile Health IT system to improve healthcare delivery in Windhoek rural health centres

A mobile health IT system was designed to improve the healthcare delivery service in Windhoek rural health centres, in Namibia [26]. The aim was to help ensure the provision of efficient and effective healthcare services to patients and to enable efficient work processes of both nurses and doctors [26]. Acceptance, perceived usefulness and ease of use of the MHSF was crucial in this study and UCD provides an approach for evaluating these factors. As previously mentioned, the UCD approach begins with a thorough understanding of the needs and requirements of the user. Interaction amongst potential users was very important for the MHSF and based on the UCD approach, establishing the context in which users may use the system should be defined at the beginning. Therefore, participants involved in requirement elicitation were purposely selected. This involved conducting the study in rural areas and early integration of potential users involved in the delivery and receiving healthcare delivery service in the design process.

AT helped provide a structure and a richer understanding of the needs of subjects/users as well as their related activities/tasks. This study demonstrates an emphasis on using AT in investigating the patient needs and requirements; activities are then separated into subjects, tools and objects. The healthcare delivery needs and requirements in Windhoek rural hospitals were then analysed using constructs from AT within a UCD framework, which also included evaluating acceptance of the mobile phones using principles of UCD [26]. Reflecting on Engeström's activity system model (Figure 3), we note the different constructs of AT and how these are featured in this case study. AT is used to understand the interaction amongst the subjects (doctors, nurses and patients) and the objects (activities and processes involved in providing and receiving healthcare delivery service). The tools in this study are the patient health cards used to record patient health information and activities; this was described as a mediating tool between the patients, nurses and doctors. The rules guiding these activities include the queuing and payment process, vital signs checking, diagnosis and drug prescription. The community which takes part in these activities include the doctor, the nurses and patients, and within these activities, work is divided among them. The nurses are responsible for checking vital signs; the doctors are responsible for diagnosing the patients and prescribing drugs while the patients are receiving this healthcare delivery service.

Structured interviews were conducted with doctors, nurses and patients to investigate the healthcare needs of patients. The findings helped establish the current work process, daily activities in the rural health centres, needs and requirements of the patients. These subsequently informed the design and development of the MHSF. A doctor, a nurse, IT specialists and researcher then provided expert opinion on the proposed MHSF. Although there was no iteration in this study, it was indicated that the framework would be expanded which would lead to another iteration of the UCD approach. Hence some of the requirements might change or be redefined [26].

Table 2 illustrates how UCD principles were enabled by applying activity theory. The table highlights the different stages during the UCD lifecycle and how the related UCD principles are applied to the relevant stages in the featured Use Case by using AT within UCD methods.

Table 2. Summary of the application of UCD and AT in the requirements, evaluation stages of the Mobile Health Service Framework.

UCD Lifecycle Stage	UCD Principle	Application of UCD and AT in the Use Case
Understand and specify the context of use	Gain a clear understanding of users, healthcare tasks and environment.	Analysing the characteristics of the users and relevant stakeholders, their health activities, work processes and the environmental conditions in which the system will be used. Structured interviews were conducted with a group of potential users to identify the healthcare needs of patients. Applying UCD helped with the definition of the users and to understand the context of use. Users here are doctors, nurses and patients. UCD also facilitated understanding the needs of potential users early in the requirement process. AT helped identify and understand the interaction amongst the users, the activities they perform and the use of the patient health card as a mediating tool. Findings from the study were also analysed using constructs of AT. Two criteria were considered for acceptance, perceived usefulness and perceived ease of use.
Establish the user and/ or organisational requirements	Formal specification of user requirements, inferred from the defined context of use.	The requirements are based on both the needs of patients, the work process of doctors and nurses in the healthcare delivery service. The requirements of the MHSF include reduced waiting times and appointment durations, and SMS facilities regarding health information.
Produce design solutions	The design solutions are produced while trying to meet user requirements as much as possible.	Transforming user needs and requirements to inform the development of the MHSF.
Evaluate design against requirements	The design is evaluated.	The proposed MHSF was evaluated using an expert review on the acceptance, usability, perceived usefulness and ease of use.

2.3. Use Case 2: Implementation evaluation of a mHealth system used by community health workers

AT can also provide a theoretical lens for adopting a UCD approach to evaluating systems in the deployment phase. MomConnect is a mHealth system used by community health workers in South Africa to provide advice to pregnant women via SMS [27]. The system was part of a government initiative to improve public services due to the high South African rates of pregnant women facing poverty and multiple health conditions. The users of the system are the community health workers, the pregnant women and the clinic managers. Community health workers use the system to register pregnant women, to enable them to receive messages from the system. Pregnant women use the system to receive advice and appointment reminders via SMS. Clinic managers run the antenatal clinics where registration takes place and have access to the system data [27].

AT was used to help understand usage of this mHealth system from the perspective of the users and to provide analysis on the key drivers of use. This study was able to bridge the gap between the limited understanding of how innovations are adopted in practice, and how this relates to specific characteristics of the technical system and users. AT was used to study routine use of the system by the observation of subjects (community health workers, pregnant women and clinic managers) and the processes of registration activities, as well as the connection between the tools and the community involved in these activities. The analysed data were used to provide a broader understanding of user needs and to implement the requirements into an improved contextual technology fit. This use case also illustrates the reiterative approach applied, in accordance with a UCD framework, commencing with an evaluation of the system in deployment and then using that analysis to inform the requirements of a system redesign (see Table 3).

Table 3. Application of UCD and AT in evaluating the implementation of *MomConnect.*

UCD Lifecycle Stage	UCD Principle	Application of UCD and AT in the Use Case
Understand and specify the context of use	Gain a clear understanding of users, healthcare tasks and environment.	This stage consists of examining the user context. Drawing on the features of UCD, potential users were collaborated with throughout the process. It helped to establish who the users of the system are and to understand the context of use. AT facilitated the breaking down and binding of entities with design tasks and goals. It facilitated the understanding of the influence of the mHealth solution in the context of existing work practices, tools and the broader context of the health facility.
Establish the user and/ or organisational requirements	Formal specification of user requirements, inferred from the defined context of use.	This included gathering the information that informs the design of the system. Open-ended interviews and observation were conducted with community health workers and clinic managers to understand key drivers and constraints of use. AT was used as a framework to provide a structured set of concepts used to analyse the goal-oriented action. AT was used to analyse the dynamics of mHealth use by the facility staff.
Produce design solutions	The design solutions are produced while trying to meet user requirements as much as possible.	The findings were used to inform the proposed concept of affordance that is focused on distinguishing between the situation where user groups and work practices in facilities are centrally or peripherally targeted by the designers. It was also used to contribute to theory more broadly by developing an Activity Theory-framed approach to affordance actualisation.
Evaluate design against requirements	The design is evaluated.	UCD facilitated a continuous evaluation of the system following deployment. In this case, after an initial round of data collection, the researcher returned to the field some months later, to conduct additional observation sessions so that they could develop a fuller picture of how the usage practices were playing out.

3. Explanation of success or failure of mHealth systems

It has previously been emphasised within this chapter that there is a fundamental need to ensure that mHealth interventions demonstrate strong theoretical underpinnings to ensure efficacy, effectiveness and acceptance, as well as other important factors relating to optimum usage. This section firstly provides a brief overview of critical success factors used in mHealth. It then examines the determinant critical success or failure factors that relate to the aforementioned use cases, and how an HCI approach that links AT as a theoretical lens to a UCD framework can help support these factors.

Research shows that the: quality; intention to use; efficiency; usability; trust and increased user satisfaction are critical success factors for healthcare systems [28,29,30,31]. However, successful interactive technologies are not simply usable, rather they should provide engaging experiences that are highly sensitive to the use context, particularly the expectations, goals, motivations, and needs to be possessed by their users [32]. Perceived value and ease-of-use are critical factors in the successful adoption of a mHealth system that is used to self-manage health conditions [29]. It is then critical that the mHealth intervention is simple, intuitive and achieves its goal of enabling users to improve self-management of the health condition [33]. End-user involvement in the design and implementation of mHealth systems is an important determinant to the eventual success and for enabling optimum clinical impact [34]. As previously discussed, involving end users throughout the development lifecycle brings new insights for customising the technology to provide a better fit to requirements. Trust in relation to users' concern over the security of personal data can influence the intention to use [28]. Low acceptance, adoption, end-user levels of technical literacy are also barriers and can impact upon the efficacy of the mHealth system [29]. Lastly, negative perceptions toward mHealth systems can significantly reduce users' willingness to adopt new technology [29].

Tables 4-5 summarise the critical success factors in each use case.

Table 4. The determinant factors for success using AT and UCD in Use Case 1.

Critical Success Factor	Description within the use case: *MHSF*
Ease-of-use	Users rated usability in relation to perceived usefulness and perceived ease-of-use scales and were then able to provide richer feedback via open-ended interviews. AT facilitated an 'activity-oriented interactive flow' to the design of the system, which users commented helped contribute to the ease-of-use of the system. The iterative feature of UCD facilitated feedback from users that could be implemented for the next design iteration of *MHSF*. Given that the system may be required to monitor more serious health conditions, it is essential that the system is designed to ensure ease-of-use to reduce the complexity of user managing these conditions.
User satisfaction	Using AT provided a richer analysis of user needs and activities, which helped in designing a system that affords increased user satisfaction. MHSF was evaluated by potential end users to ensure that user needs were met. Ease-of-use and user satisfaction are strongly connected. Where optimum ease-of-use is enabled, users are more likely to be satisfied with the system [34]. User satisfaction was also measured during the evaluation of the system design using interviews informed by the constructs AT.

Motivation & acceptance	Using the AT and UCD approach helps to increase the likelihood of user motivation and acceptance by collaborative decision making and to place the users at the forefront of the design. The evaluation of *MHSF* enabled users to provide feedback on the acceptance. Acceptance was analysed using constructs that includes perceived usefulness and perceived ease of use.
Trust	Trust is critical to the success of a mHealth system [28]. Evaluating user perspectives of the system during interviews highlighted that users were concerned about the security of their personal information. These concerns can then be implemented and evaluated in the next iteration of the project design.
Confidence in use	AT coupled with UCD enabled a broad understanding of user needs which also includes assessing aspects of technical literacy. The UCD approach helped to understand how well users who lack technical literacy were able to use the system. MHSF was designed to be intuitive with good ease-of-use to enable usage by people with limited computer proficiency. Intuitive and user-friendly designs could help to increase user confidence.

Table 5. The determinant factors for success using AT and UCD in Use Case 2.

Critical Success Factor	Description within the use case: *MomConnect*
Efficiency	Using AT in an observational ethnographic study, coupled with the interview, the results from the deployment evaluation study indicated that the registration process on the *MomConnect* system did not afford optimum efficiency, leading to complaints from pregnant women to the clinic because they had not been registered. By taking the perspective and practices of the staff, AT helped to describe how *MomConnect* can be used in the clinic, by presenting a series of nested activity system with different goals. AT helped break down goals to improve the operation of the clinic to enable effective registration on *MomConnect*. Breaking down the goals will then inform the redesign of the *MomConnect* system by allocating resources in a way that promotes efficiency.
Motivation & acceptance	Using AT to evaluate the system in use, interviews highlighted that acceptance was driven by motivation from not only the main users but clinic managers themselves. The analysis showed that these clinic managers who are not involved in the registration process, do in fact have some influence over the practice of using *MomConnect* to register pregnant women. The study demonstrated that motivation to use the *MomConnect* could be influenced by the level and type of clinical management. Women would use the system if it were recommended and benefits are promoted by more senior members of staff because they are more likely to be assigned as 'experts'.
Trust	A UCD approach to evaluating the system in use enabled users to identify confidentiality concerns during the interviews and ethnographic observational studies. This was important to address in the redesign of MomConnect to help increase the future success of the system. Users' perception of the confidentiality of their data, needs to be considered in context to inform the next iteration of UCD design, the requirement stage.
Confidence in use	The ethnographic observational studies, as well as interviews, helped identify issues relating to technical literacy in both staff and the pregnant women. Lack of technical literacy will impact upon user's confidence. This analysis will help in the design of a more intuitive system in future iterations.

4. Discussion

mHealth technologies must be designed to meet the wide spectrum of end-user needs, as well as enable optimum acceptance and clinical impact, where relevant. There are several examples of how the interdisciplinary field of HCI has been instrumental in providing the necessary tools in designing these technologies [7,15,35,36]. The application of HCI in mHealth has seen the employment of multidisciplinary theoretical frameworks including activity theory, as well as others including distributed cognition and cognitive ergonomics. HCI has also given rise to the development of design frameworks such as UCD which makes use of several user centred methods used within the reiterative project life cycle. UCD has been shown to be an effective framework in the design of mHealth interventions by linking Activity Theory to provide a theoretical lens during various stages of the project lifecycle [26,27]. Activity theory can help provide a broader framework for understanding human computer interaction. The elements within Engestrom's model of activity help to provide a multifaceted analysis of users, their activities and the relationships between them. The use cases presented in this chapter illustrate how AT combined with UCD can be applied throughout the different stages of the intervention lifecycle for the mobile intervention, from analysis through to deployment. The use cases also illustrate how the application of AT and UCD can help maintain or solve success or failure factors of mHealth systems. Whilst the benefits of using this approach have been illustrated, some discussion around its limitations can help in providing a critically balanced argument for its use in mHealth.

Whilst AT can provide valuable insights into understanding user needs and their activities, the theory itself can be difficult to comprehend, particularly for system designers that are not from a cultural-historical/psychology background. It is not a rigid theory and does require some understanding of its historical context to be able to utilise its principles in practice. Engestrom's model of activity (Figure 3) highlights the complexity of understanding human activities however, it can be difficult for those involved with system design to decompose the model to specific focal points in the design [36]. AT can also be a time-consuming process and care needs to be taken to ensure that this process does not impact negatively upon time constrained subjects.

There are however many examples in healthcare and other domains, where it has been advantageous to complement AT with other methods. For example, combining AT with cognitive load theory and flow experience theory to enable the development of a more integrated framework for analysing internet-mediated experiences of children, as well many examples of AT combined with distributed cognition theory to provide a conceptual framework for Computer Supported Collaborative Work (CSCW) research [37]. The application of AT does enable a broader understanding of conceptualising human in context, which is particularly relevant when trying to understanding how humans will interact with systems. The importance of undertaking a comprehensive user analysis cannot be understated. If users' needs and expectations are not met, then this will inevitably impact upon the success of a system. mHealth systems that are designed for users to help manage their healthcare, not only need to be usable but must also enable clinical impact where relevant. Acceptance of the intervention and providing a positive user experience are then key. Linking AT to a UCD framework that involves users throughout the project lifecycle of a mHealth system can help to achieve these goals.

Teaching questions for reflection

1. How can AT be used as a conceptual framework in UCD to gain a better understanding of user needs in the design of mHealth healthcare systems?
2. How can we increase the likelihood of user acceptance of mHealth systems?
3. Critically evaluate the strengths and weaknesses of using Activity Theory as a theoretical lens in a UCD approach.
4. What are some of the critical success factors of mHealth interventions, and how can an AT and UCD approach help?

References

[1] Z. Wang, L. Zhang, L. Ma, N. Liu. Modeling Medical Services with mHealth Applications, *Journal of Healthcare Engineering*, **61** (2018), 1538-1554.
[2] E. Poole. HCI and mHealth interventions: How human–computer interaction can contribute to successful mHealth interventions, *Translational Behavioral Medicine,* **3** (2013), 402–405.
[3] F. Iacobelli, R. Adler, D. Buitrago, J. Buscemi J, Corden. Designing an mHealth application to bridge health disparities in Latina breast cancer survivors: a community-supported design approach, *Design for Health*, **2** (2018), 58-76.
[4] C. Abras, D. Maloney-Krichmar, J. Preece. User-Centered Design. In: Bainbridge W, editor. *Encyclopedia of Human-Computer Interaction*. Sage, Thousand Oaks, 37 (2004), 445-456.
[5] D. Vollmerdahlke, K. Fair, Y. Hong. Apps Seeking Theories: Results of a Study on the Use of Health Behavior Change Theories in Cancer Survivorship Mobile Apps, *JMIR mHealth and uHealth,* **3** (2015), p.e31.
[6] S. Cigdem. Rules of engagement in mHealth: what does mHealth bring to research and theory? *Contemporary Nurse*, **54** (2018), 374-387.
[7] T. Hewett, R. Baecker, S. Card, T. Carey, J. Gasen, M. Mantei. *ACM SIGCHI curricula for human-computer interaction*. New York, NY: Association for Computing Machinery, (1995).
[8] A. Teixeira F. Ferreira, N. Almeida, S. Silva. Design and development of Medication Assistant: older adults centred design to go beyond simple medication reminders, *Universal Access in the Information Society*, **16** (2017), 545-560.
[9] W. Hick. On the rate of gain of information, *Quarterly Journal of Experimental Psychology,* **4** (1952), 11–26.
[10] P. Fitts. The information capacity of the human motor system in controlling the amplitude of movement, *Journal of Experimental Psychology,* **47** (1955), 381–391
[11] D. Norman. *The Design of Everyday Things*, New York, NY: Basic Books (2013)
[12] L. Vygotsky. Mind in society: *The development of higher psychological processes*, Cambridge, MA: Harvard University Press, (1978).
[13] S. Card, T. Moran, A. Newell. *The Psychology of Human Computer Interaction*, New Jersey, Lawrence Erlbaum Associates, (1983).
[14] S. Card, T. Moran, A. Newell. The keystroke-level model for user performance time with interactive systems, *Communications of the ACM*, **23** (1980), pp 396–410.
[15] G. Kim. *Human–Computer Interaction: Fundamentals and Practice*. Auerbach Publications, 2015.
[16] L. Uden, P. Helo. Designing mobile interfaces using activity theory, *International Journal of Mobile Communications*, **5** (2008), 616-32.
[17] A. De Vito Dabbs, B. Myers, K. Mc Curry, J. Dunbar-Jacob. User-centered design and interactive health technologies for patients, *Computers, informatics, nursing: CIN*, **27** (2009), 175-83.
[18] E. Baek, K. Cagiltay, E. Boling, T. Frick. User-centered design and development. *Handbook of research on educational communications and technology*, **1** (2008), pp.660-668
[19] ISO. ISO 9241-210:2010 *Ergonomics of human-system interaction -- Part 210: Human centred design for interactive systems*, 2010.
[20] C. Johnson, T. Johnson, J. Zhang. A user-centered framework for redesigning health care interfaces, *Journal of biomedical informatics*, **1** (2005), 75-87.
[21] J. Roy, W. Neumann, D. Fels. User Centered Design Methods and Their Application in Older Adult Community. *International Conference on Human Interface and the Management of Information*, (2016), 462-472.

[22] D. Mwanza. *Where theory meets practice: A case for an activity theory based methodology to guide computer system design*, Knowledge Media Institute, The Open University, Milton Keynes. UK. (2001).

[23] Y. Engeström. *Learning by Expanding: an Activity-theoretical Approach to Developmental Research*. Helsinki: Orienta-Konsultit Oy, (1987).

[24] M. Maguire, N. Bevan. User requirements analysis, *IFIP World Computer Congress*, 13 (2002), 133-148.

[25] D. Hoang, D. Chiang Choon Poo. Analysis and design of mHealth interventions towards informed shared decision making: an activity theory-driven perspective, *Journal of Decision Systems*, **25** (2016), 397-409.

[26] A. Coleman, G. Iyawa. Improving health care delivery in rural communities through the use of mobile phones: a case study in Windhoek, *International Science and Technology Journal of Namibia*, **6** (2015).

[27] B. Wolff-Piggott, U. Rivett. An Activity Theory Approach to Affordance Actualisation in mHealth: the Case of MomConnect, *European Conference on Information Systems*, (2016), 108

[28] D. Sensuse, P. Kareen, H. Noprisson, M. Pratama. Success factors for health information system development. *IEEE International Conference on Information Technology Systems and Innovation*, (2017) 162-167.

[29] N. Matthew-Maich, L. Harris, J. Ploeg, M. Markle-Reid, R. Valaitis, S. Ibrahim, A. Gafni, S. Isaacs. Designing, implementing, and evaluating mHealth technologies for managing chronic conditions in older adults: a scoping review, *JMIR mHealth and uHealth* **4** 2016.

[29] C. Johnson, T. Johnson, J. Zhang. A user-centered framework for redesigning health care interfaces, *Journal of biomedical informatics*, **38** (2005), 75-87.

[30] S. Van der Weegen, R. Verwey, M. Spreeuwenberg. The development of a mobile monitoring and feedback tool to stimulate physical activity of people with a chronic disease in primary care: a user-centered design, *JMIR mHealth and uHealth*, **1** (2013), p.e8

[31] D. Organ, S. Dick, C. Hurley. A systematic review of user-centred design practices in illicit substance use interventions for higher education students, *European Conference on Information Systems: Beyond Digitization-Facets of Socio-Technical Change,* AIS Electronic Library 2018.

[32] J. Nielsen, S. Mengiste. Analysing the diffusion and adoption of mobile IT across social worlds, *Health Informatics Journal*, **20** (2014), 87-103.

[33] J. Nielsen, J. Levy. Measuring Usability: Preference vs. Performance. *Commun. ACM.* **37** (1995), 66-75.

[34] D. Shellmer, M. Dew, G. Mazariegos, A. DeVito Dabbs. Development and field testing of Teen Pocket PATH®, a mobile health application to improve medication adherence in adolescent solid organ recipients, *Pediatr Transplant*, **20** (2016), 130–140.

[35] P. Collins, S. Shukla, D. Redmiles. Activity theory and system design: A view from the trenches. *Computer Supported Cooperative Work*, **11** (2002), 55-80.

[36] C. Torkil, V. Kaptelinin, B. Nardi. Making HCI theory work: an analysis of the use of activity theory in HCI research, *Behaviour & Information Technology*, **35** (2016), 608-62.

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

doi:10.3233/SHTI190111

Technology Acceptance Models in Health Informatics: TAM and UTAUT

Elske AMMENWERTH[a,1]

[a] *UMIT – University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria*

Abstract. Both the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT) aim at understanding better why users accept or reject a given technology, and how user acceptance can be improved through technology design. Two case studies are presented where TAM and UTAUT were successfully used in a health care setting to predict technology adoption. Both models have found popularity in health care. However, recent reviews show that TAM and UTAUT failed to provide stable predictive capabilities for acceptance and use of technologies in health care. Reasons for this may be the specific context of health care, where not only the technology, but also socio-organizational and cultural factors influence technology acceptance.

Keywords. Health informatics; Assessment, technology; Intention; Attitude to computers; Models, theoretical

Learning objectives

After reading this chapter the reader will understand:

1. How TAM and UTAUT attempt to predict and explain technology acceptance and technology usage.
2. How TAM and UTAUT can be applied in health informatics projects to predict and support health IT acceptance and usage.
3. The strengths and limitations of TAM and UTAUT with regard to prediction of the support needed for health IT acceptance and usage.

1. Introduction to TAM and UTAUT

This section introduces TAM and UTAUT as two technology acceptance models. Both have several similarities, which is not surprising, as UTAUT was among others developed based on TAM. We decided to focus on TAM and UTAUT, from a longer list of available technology acceptance theories (for an overview, see for example [1, 2]), as both have found widespread adoption in health care.

[1] Corresponding Author: Elske Ammenwerth, E-mail: elske.ammenwerth@umit.at.

Both TAM and UTAUT aim at predicting technology usage by looking at the factors that will influence technology acceptance. Both theories focus on two questions to explain technology acceptance and resulting technology use: Is the technology useful for me? And: Is the technology easy to use?

UTAUT adds two further questions to this list: Does my social environment want me to use the technology? And: Do I have the necessary technical and organizational infrastructure to use the technology?

Now let's have a more detailed look on TAM and UTAUT.

1.1. Technology Acceptance Model (TAM)

User acceptance is often a pivotal factor in the success or failure of a new information system [3]. The goal of the Technology Acceptance Model (TAM) is to understand better why users accept or reject a given technology, and how user acceptance can be improved through technology design. TAM was developed by Fred D. Davis in the late 1980s [4, 5].

The Technology Acceptance Model is based on principles from Fishbein and Ajzen's Theory of Reasoned Action [6]. TAM hypothesizes that two particular beliefs, Perceived Usefulness and Perceived Ease of Use, are of primary relevance for technology acceptance [7]. Perceived Usefulness is the expectation of a user that the system will be useful for the job. Perceive Ease of Use is the expectation that the system is user friendly and easy to use. Perceived usefulness is influenced by Perceived Ease of Use, as users will find easy-to-use systems more useful [3]. Both beliefs are determinants for Attitude towards Using. This Attitude towards Using is then a determinant of the Behavioral Intention to Use, which can be interpreted as technology acceptance [8]. The actual system usage is then determined by this Behavioral Intention to Use.

In a review of TAM usage in health care, Holden found that TAM was able to predict 30 - 70 % of variance of Behavioral Intention to Use, which can be considered reasonably high.

Table 1 summarizes the basic definition of the concept used in TAM. Figure 1 shows the TAM model.

Table 1. Definitions of the variables used in TAM [8].

Perceived Usefulness	An individual's perception that using an IT system will enhance job performance.
Perceived Ease of Use	An individual's perception that using an IT system will be free of effort.
Attitude toward Using	An individual's evaluative judgment of the target behavior on some dimension (e.g., good/bad, harmful/beneficial, pleasant/unpleasant).
Behavioral Intention	An individual's motivation or willingness to exert effort to perform the target behavior.
Use	One specific behavior of interest performed by individuals with regard to some information technology (IT) system.

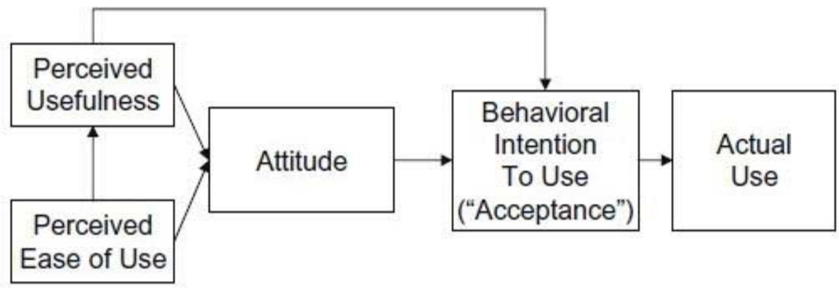


Figure 1. Technology Acceptance Model (TAM) [8].

A number of TAM extensions have been proposed to overcome some limitations in the original model. Several studies added single variables to the original TAM to increase the predictive power of the model, such as individual user factors, organizational readiness, or trust - an overview is given by Holden [8]. TAM2 by Venkatesh & Davis (2000) [9] extended TAM with variables that were seen as influencing perceived usefulness or user acceptance, such as subjective norm, image, voluntariness of use, or job relevance. Further extensions of TAM, such as TAM 3 [10], added other concepts such as computer anxiety or enjoyment.

The TAM theory of technology acceptance and use has gained significant popularity in the field of technology acceptance research and is considered a "key model" [11] or "gold standard" [8] in understanding predictors for IT acceptance. A PubMed query on "Technology acceptance model" retrieved 340 papers (search done on 7 June 2018), showing the popularity of TAM also within health informatics research.

1.2. Unified Theory of Acceptance and Use of Technology (UTAUT)

The Unified Theory of Acceptance and Use of Technology (UTAUT) was published by Venkatesh and Davis in 2003 [2]. UTAUT is based on an analysis and comparison of eight technology acceptance models, among them TAM, TAM2, the Theory of Reasoned Action and the Diffusion of Innovation Theory. The aim was to synthesize the multitude of available models on technology acceptance into one unified model. The aim of UTAUT is to assess the likelihood of success for new technologies and to understand drivers of acceptance [2].

UTAUT describes four key variables: Performance Expectancy, Effort Expectancy, Social Influence, and Facilitating Conditions. Performance Expectancy is the expectation of a user that the system will be useful for the job; it corresponds to Perceived Usefulness in TAM. Effort Expectancy is the expectation that the system is user friendly and easy to use; it corresponds to Perceived Ease of Use in TAM. Social influence is defined as the degree to which a user perceives that important others believe he or she should use the new system. Facilitating Conditions are defined as the degree to which a user believes that an organizational and technical infrastructure exists to support system use.

In UTAUT, Behavioral Intention to Use the technology is determined by Performance Expectancy, Effort Expectancy, and Social Influence. Actual Usage is then determined by Behavioral Intention to Use and the Facilitating Conditions. Gender, age, experience, and voluntariness of use moderate the impact of the key variables on usage intention and behavior.

The original UTAUT validation study [2] found that UTAUT was able to explain 70% of the variance of Behavioral Intention, which indicates high predictive power.

Table 2 summarizes the basic definition of the variables used in UTAUT. Figure 2 shows the UTAUT model. The similarities to TAM are obvious when comparing both models.

Table 2. Definitions of the variables used in UTAUT [2].

Performance Expectancy	Degree to which an individual believes that using the system will help him or her to attain gains in job performance.
Effort Expectancy	Degree of ease associated with the use of the system.
Social Influence	Degree to which an individual perceives that important others believe he or she should use the new system.
Facilitating Conditions	Degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system.
Behavioral Intention to Use	Measure of the strength of one's intention to perform a specified behavior.

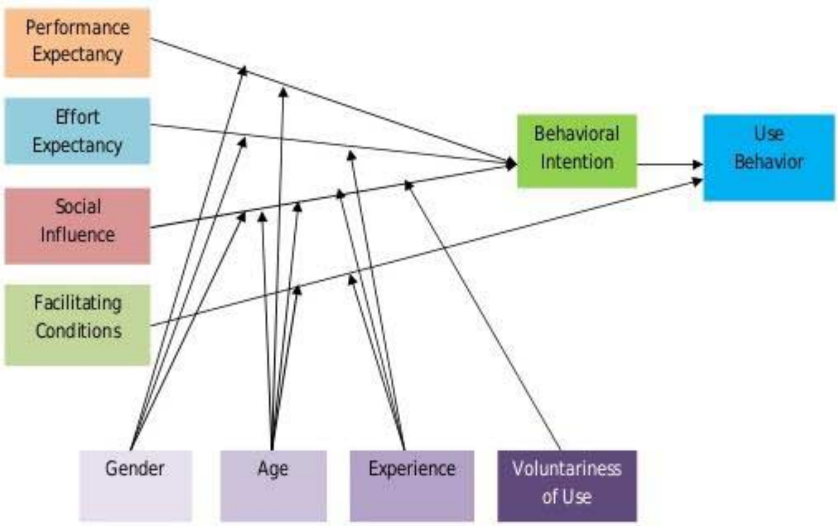


Figure 2. Unified Theory of Acceptance and Use of Technology (UTAUT) [2].

A PubMed query on "Unified Theory of Acceptance and Use of Technology" retrieved 80 papers from 2008 onwards (search done on 7 June 2018), showing the popularity of UTAUT within health informatics research.

2. Usage of TAM and UTAUT in health informatics

This section describes two use cases of health informatics interventions where TAM and UTAUT were applied.

2.1. Case Study 1: Perceived usefulness and perceived ease of use of electronic health records among nurses: Application of Technology Acceptance Model

In this study of Ahmad Tubaishat from 2017 [12], 1.539 nurses from 15 Jordanian hospitals using a nursing electronic health records (EHR) were surveyed using a 28-item questionnaire based on TAM.

Results show that the nurses demonstrated a positive perception of the usefulness and ease-of-use of EHRs, and their acceptance of the technology. Both Perceived Usefulness and Perceived Ease of Use had an influence on the intention to use EHRs: The effect of Perceived Usefulness explained 51% of the variance of intention to use EHRs, whereas Perceived Ease of Use predicted 42% of the variance.

Perceived Usefulness was affected by gender, professional rank, EHR experience, and computer skills of the nurses, these variables explained 55% of the variance of Perceived Usefulness. The Perceived Ease of Use was affected by nursing experience, EHR experience, and computers skills, these variables explained 44% of the variance of Perceived Ease of Use.

The authors concluded that training should include also basic computer skills, as this may positively influence Perceived Ease of Use and Perceived Usefulness and by this may increase EHR acceptance.

The case study shows how the TAM based survey can help to identify factors for further optimization of health IT implementation. In particular, it shows which variables influence directly or indirectly Behavioral Intention to Use in this context, opening ways to intervene e.g. by training.

2.2. Case study 2: Analyzing older users' home telehealth services acceptance behavior, applying an extended UTAUT model

In this study by Miha Cimperman from 2016 [13], 400 Slovenian participants aged 50 years or older were surveyed on their acceptance of a home telehealth service using UTAUT. Respondents were randomly selected equally across all regions. Respondents first got a short written explanation of the idea of home telehealth service and were then asked on their opinion on this. The survey comprised 47 standardized items based on UTAUT. Three context specific predictors were added to the original UTAUT model, namely Doctor's Opinion, Computer Anxiety, and Perceived Security.

As expected, Performance Expectancy, Effort Expectancy, Facilitating Conditions, and Perceived Security were found to have a direct impact on Behavioral Intention to Use the home telehealth service. In addition, Computer Anxiety was found to be an antecedent of Effort Expectancy with a strong negative influence, and Doctor's Opinion influence showed a strong positive impact on Performance Expectancy. Different to the UTAUT assumption, Social Influence was not a predictor of Behavioral Intention, which authors explained by the fact the elderly users may not be so much dependent on social pressure [13]. The model of the six predictors explained 77% of the total variance of Behavioral Intention to Use, indicating a strong predictive power of the revised model [13].

The authors concluded that health professionals should be involved as social agents to frame home telehealth services as useful and beneficial, as this will raise acceptance among the users. Also, home telehealth services should be promoted as secure, to build trust. Different types of technical equipment should be made available to reduce computer anxiety. Due to low social influence, they see it as unlikely that early adopters as pioneer users will contribute significantly to the diffusion of home telehealth services among other users.

The case study shows how the UTAUT-based survey can help to identify factors for the further optimization of health IT implementation before the technology is introduced. This case study is one of the few examples where TAM or UTAUT were used in settings where the technology was yet to be implemented.

3. Explanation of success or failure of health IT system

TAM and UTAUT have been developed with the aim to understand better why users accept or reject technology, and to predict acceptance or non-acceptance of new technology. TAM and UTAUT define acceptance as the intention to use, or the willingness to use, a technology. The theories assume that intention to use is a direct determinant for actual system use. Thus, TAM and UTAUT attempt to reveal factors that have direct implication for the success or failure of technology, with success seen as equivalent to actual system usage.

We must note that TAM und UTAUT have not been developed within a health care setting. TAM was developed based on studies of an e-mail system and a word processing system [3]. UTAUT was validated based on studies related to introducing an online meeting manager, a database application, and an accounting system [2]. These types of application seem not comparable to much more complex health care technologies, such as computerized physician order entry systems, electronic health record systems, or nursing documentation systems. In addition, these latter types of technologies represent socio-technical information systems where the acceptance of a technology depends not only on its functionality or ease of use, but on many other factors such as hardware performance, training, support, and workflow integration. In particular, besides individual factors shaping decisions to use a technology, organizational, cultural and emotional factors will also influence technology acceptance in healthcare settings [14]. Overall, these socio-organizational-cultural factors are not well covered by TAM and UTAUT. Besides, TAM was developed with a focus on technology which can be used voluntarily. Typically, in health care, most technologies are mandatory to be used by the staff. This all distinguishes healthcare from the settings where TAM and UTAUT were developed and used.

Nevertheless, as the case studies and a short query in PubMed show, TAM and UTAUT have found wide adoption in health care. Reasons for this can be the quite simple assumptions of both models: System usage depends on only two (TAM) or four (UTAUT) key variables, including the usefulness of the system for the work and the ease-of-use of the system. This sounds quite intuitive and may have contributed to their popularity.

Still, in many health care studies where TAM or UTAUT were applied, authors have added variables to extend the original TAM or UTAUT models to better adapt it to the context of health care. Case study 2 [13] showed an example of this: The authors assess the acceptance of home telehealth services by elderly patients and added three context-specific predictors, namely Doctor’s Opinion, Computer Anxiety, and Perceived Security. The authors argue that the “universal” variables in UTAUT are not specific enough for health care and thus decided to add these three “context-specific” variables as potentially important predictors for the acceptance of the telehealth service. And indeed, all three context variables were found to be important predictors in the study.

Holden [8] lists several other examples where TAM or UTAUT were extended by context-specific variables. This shows that despite their large popularity, both models may need to be parsimoniously applied in more complex health care settings.

Holden also points to the fact the key variables of TAM and UTAUT are not measured uniformly in different studies. Instead, studies often modify original survey items to adapt the questions to the local study context (either by rewording questions or by adding completely new questions). All this shows TAM and UTAUT are somewhat unspecific for health care settings.

In general, both TAM und UTAUT have been found to predict Intention to Use quite well, with explained variance up to 70%. Yet, closer analysis to their application in health care by Holden (2010) shows that only Perceived Usefulness was consistently found to be a significant predictor of Intention to Use (in all of the 16 reviewed studies) [8]. In contrast to this, Perceived Ease of Use was found to be a significant predictor of Intention to Use in seven of 13 studies only [8]. And Social Influence, an UTAUT variable, was found to be significant predictor in four of eight studies only. Also Gücin (2015) states, based on a literature review, that Perceived Usefulness is "the most powerful predictor of the technology acceptance" [15].

Summarizing these findings, we see that the key assumptions of TAM and UTAUT could not be confirmed in a large number of technology acceptance studies in health care. These findings indicate that health care is indeed a special setting where the simple assumption of TAM and UTAUT may not fully match the more complex reality. Holden (2010), for example, summarizes that Perceived Ease of Use may not be that important for technology acceptance and usage when users are sufficiently experienced with the system or when they have sufficient IT support. Also, Social Influence may not influence physicians as users so strongly, as they are more independent and "immune to peer pressure" [8]. Also, after an analysis of several acceptance theories, Gücin (2015) points to the fact that the acceptance factors for health care professionals and patients may be different, with patients seeing for example ease of use as more important than health care professionals [15]. Also, he argues that important acceptance factors such as suspicions of confidentiality and privacy or individual characteristics of the user (e.g. of early adopters) may be strong influencing factors, but these are not considered in the original models [15].

To conclude, while TAM and UTAUT have been broadly adopted as a means of predicting technology acceptance and usage, the findings in health care are quite mixed. Both the fact that many studies in health care cannot find support for some basic hypothesis of TAM and UTAUT, and the fact that many authors added variables to the original TAM and UTAUT models or revised the survey instruments to respond to context influencing factors, point to the fact that the original TAM and UTAUT fail to demonstrate strong predictive capabilities for technology acceptance in health care [14].

4. Conclusion

TAM and partly UTAUT provide a more technology-centered view on technology acceptance, where acceptance is understood to mostly depend on the nature of technology [14], i.e. functionality and ease of use. Socio-organizational, workflow, cultural or emotional aspects as well as differences in user groups (physicians, nurses, patents) are not well covered [14] and may explain why in several studies in health care, basic assumptions of the model could not be confirmed.

Overall, when applied to health care settings, TAM and UTAUT failed to provide stable predictive capabilities for technology acceptance and use[2].

Teaching questions for reflection

1. What are the major differences between TAM and UTAUT?
2. How do TAM and UTAUT accommodate for socio-organizational or cultural factors for technology acceptance and technology usage?
3. How could you use TAM or UTAUT when preparing for the hospital-wide introduction of a nursing documentation system?

References

[1] R. Sharma, R. Mishra, A review of evaluation of theories and models of technology adoption, *Indore Management Journal*, **6**(2) (2014),17-29.

[2] V. Venkatesh, M. Morris, G. Davis, F. Davis, User Acceptance of Information Technology: Toward a Unified View, *MIS Quarterly*, **27**(3) (2003),425-78.

[3] F. Davis, User acceptance of information technology: System characteristics, user perceptions and behavioral impacts, *International Journal of Man-Machine Studies*, **38**(1993),475-87.

[4] F. Davis, A technology acceptance model for empirically testing new end-user information systems - theory and results, *PhD thesis* (1985),Massachusetts Institute of Technology.

[5] F. Davis, Perceived usefulness, perceived ease of use, and user acceptance of information technology, *MIS Quarterly*, **13**(1989),319-9.

[6] I. Ajzen, M. Fisbhein, Attitude-behavior relations: a theoretical analysis and review of empirical research, *Psychological Bulletin*, **84**(1977),888-918.

[7] F. Davis, P. Bagozzi, P. Warshaw, User acceptance of computer technology - a comparison of two theoretical models, *Management Science*, **35**(8) (1989),982-1003.

[8] R. Holden, B. Karsh, The technology acceptance model: its past and its future in health care, *J Biomed Inform*, **43**(1) (2010),159-72.

[9] V. Venkatesh, F. Davis, A theoretical extension of the technology acceptance model: four longitudinal field studies, *Management Science*, **46**(2) (2000),186-204.

[10] V. Venkatesh, H. Bala, Technology Acceptance Model 3 and a Research Agenda on Interventions, *Decision Science*, **39**(2) (2008),273-312.

[11] N. Marangunić, A. Granić, Technology acceptance model: a literature review from 1986 to 2013, *Universal Access In The Information Society*, **14**(1) (2015),81-95.

[12] A. Tubaishat, Perceived usefulness and perceived ease of use of electronic health records among nurses: Application of Technology Acceptance Model, *Inform Health Soc Care*, (2017),1-11.

[13] M. Cimperman, M. Makovec Brencic, P. Trkman, Analyzing older users' home telehealth services acceptance behavior-applying an Extended UTAUT model, *Int J Med Inform*, **90**(2016),22-31.

[14] R. Ward, The application of technology acceptance and diffusion of innovation models in healthcare informatics, *Health Policy and Technology*, **2**(4) (2013),222-8.

[15] N. Gücin, Ö Berk, Technology Acceptance in Health Care: An Integrative Review of Predictive Factors and Intervention Programs, *Procedia - Social and Behavioral Sciences*, **195**(2015),1698-704.

[2] See also Chapter 16, “The NASSS Framework – a synthesis of multiple theories of technology implementation”.

Part 2

Social and Psychological Theories

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

doi:10.3233/SHTI190113

Distributed Cognition: Understanding Complex Sociotechnical Informatics

Dominic FURNISS[a,1], Sara GARFIELD [b, c], Fran HUSSON [b], Ann BLANDFORD [a] and Bryony Dean FRANKLIN [b, c]
[a] *University College London, Gower Street, London; UK*
[b] *Imperial College Healthcare NHS Trust, London; UK*
[c] *UCL School of Pharmacy, London; UK*

Abstract. Distributed cognition theory posits that our cognitive tasks are so tightly coupled to the environment that cognition extends into the environment, beyond the skin and the skull. It uses cognitive concepts to describe information processing across external representations, social networks and across different periods of time. Distributed cognition lends itself to exploring how people interact with technology in the workplace, issues to do with communication and coordination, how people's thinking extends into the environment and sociotechnical system architecture and performance more broadly. We provide an overview of early work that established distributed cognition theory, describe more recent work that facilitates its application, and outline how this theory has been used in health informatics. We present two use cases to show how distributed cognition can be used at the formative and summative stages of a project life cycle. In both cases, key determinants that influence performance of the sociotechnical system and/or the technology are identified. We argue that distributed cognition theory can have descriptive, rhetorical, inferential and application power. For evidence-based health informatics it can lead to design changes and hypotheses that can be tested.

Keywords. Distributed cognition, Sociotechnical, Informatics, DiCoT

Learning objectives

After reading this chapter the reader will be able to:

1. Summarise the theory of distributed cognition.
2. List different methods and frameworks that have been developed to facilitate the application of distributed cognition.
3. Give examples of how distributed cognition has been applied in health informatics.
4. Explain how distributed cognition can be used in formative and summative stages of a project life cycle.
5. Explain how distributed cognition can lead to design ideas and testable hypotheses.

1. Introduction of distributed cognition

Distributed cognition [1] is a theory that represents a radical departure from traditional notions of cognition focused on information processing in the brain. Distributed cognition focuses on information processing in sociotechnical systems, which we dub

[1] Corresponding Author: Dominic Furniss, E-mail: d.furniss@ucl.ac.uk

sociotechnical informatics. Sociotechnical informatics includes the propagation and transformation of information across natural and engineered computational systems at a sociotechnical level. For distributed cognition, the argument is that individuals form a tightly coupled system with their environment in such a way that they employ and exploit external structures in cognitive tasks, so the task of cognition is actually distributed [2]. For example, ambulance dispatch coordinators have cards (representing incidents) and a tray of slots (each slot representing an ambulance) so they can easily see what ambulances are assigned to what incidents and how many ambulances they have free without relying solely on their internal memory [3]. Furthermore, we can configure sociotechnical systems and design external structures to influence how information is transformed and propagated in teams of individuals. For example, in the London ambulance control room dispatch teams are organised by region so if they have an incident between two regions one team can more easily communicate with the team beside them about resource allocation [3]. Modern healthcare informatics faces increasing challenges in how information processing systems should be designed and organised, especially as systems become more distributed, interconnected and complex. Distributed cognition can help to understand complex sociotechnical informatics.

1.1. Distributed cognition: The basics

Distributed cognition was pioneered by Edwin Hutchins and colleagues in the early nineties. His book, *Cognition in the Wild*, is the seminal text in the area [1]. Distributed cognition distinguishes itself from other approaches by taking the information processing metaphor of the mind and suggesting that this should not be limited to the brain, broadening what can be considered part of the cognitive system [2]. Its unit of analysis is not the individual mind but a complex cognitive system, which is essentially sociotechnical in nature [4]. It is complex because it involves different artefacts and people, over time and physical space; it is cognitive because it is focused on information processing; and it is a system because it involves elements that interact to perform a task or achieve a goal. These defining features resonate well with complex sociotechnical informatics.

One of the earliest and best-known applications of the theory involved considering a cockpit as a complex cognitive system comprising the pilots, instruments, controls and reference materials [5]. Hutchins [5] examined the intricacies of how this system worked as the design of the tools and instruments, the way the pilots sat, and the way they communicated could influence performance. He showed that distributed cognition is not simply about offloading memory into the environment, whereby operators have extra reference material to aid recall, but that their tasks can fundamentally change depending on how cognition is distributed. For example, Hutchins described how *speed bugs* are adjusted on a speed dial to indicate safe parameters for landing speeds depending on an aircraft's weight. These do not act solely as an additional reference point in case the pilot cannot recall the figures that define the parameters for safe landing speeds: they actually provide a spatial range that the speed dial indicator should remain within, which is a very different form of interaction. Hutchins [5] includes these interactions to account for the cockpit system's memory.

Distributed cognition has also been shown in carefully constructed laboratory experiments. For example, Maglio and Kirsh [6] showed that experts sometimes make *epistemic actions* in the environment to simplify a problem space rather than trying to solve the problem in their head before acting. Also, Zhang and Norman [7] showed that participants handle the same problem differently depending on how it is represented; i.e., the physical affordances of items involved in the problem's representation can complicate, simplify or otherwise transform the problem space. In both studies the determinants of performance go beyond the individual mind to include cognitive processing across external artefacts.

Distributed cognition theory proposes that cognition can be distributed in three main ways [2]:

1. Cognitive processes may be distributed across members of a social group.
2. Cognitive processes may involve coordination between internal and external (material or environmental) structures.
3. Processes may be distributed through time in such a way that the products of earlier events can transform the nature of later events.

Essentially this approach highlights how individuals' cognitive processes extend into the environment, and how groups process information using different artefacts and structures across different spaces and over different periods of time.

1.2. Applying distributed cognition theory using DiCoT

Some commentators have criticised distributed cognition for being too unstructured for easy application, i.e. there is no 'off the shelf' methodology [8]. Cognitive ethnography is proposed as the main approach for studying cognition in the wild (e.g. interviews, surveys, observations and video and audio recording) [2]. However, cognitive ethnography is a group of techniques, which lack further structure and analytical support, i.e. there is still a big challenge for researchers to know what to look at, how to look, and how to link theory to data. To fill this gap, different methods have been developed that offer more support and instruction: the Resources Model [9]; Distributed Cognition for Teamwork (DiCoT) [10]; Determining Information flow Breakdown (DIB) [11]; and Event Analysis of Systemic Teamwork (EAST) [12]. The remainder of the chapter focuses on DiCoT, which we propose is the most developed method for understanding the details of situated interactions, rather than more exclusive focus on abstract information flows, networks and the coordination of information resources.

DiCoT [10] draws upon the structure of Contextual Design [13] to provide more analytical support. Distributed cognition's focus "has always been on whole environments: what we really do in them and how we coordinate our activity in them" [2, page 174]. Contextual design also has this focus but is not underpinned by a theoretical perspective. DiCoT extends this approach to use five interdependent models:

- Information flow model – focuses on how information is transformed and propagated in the system, taking tasks, activities and processes into account.
- Artefact model – focuses on how the design of tools, technologies and external representations influence the information processing of the system.
- Social model – focuses on the different roles people play in the system, with their different knowledge, responsibilities, skills and expertise.
- Physical layout model – focuses on how things are arranged in the physical environment and how this impacts the flow of information.
- Evolutionary model – focuses on how cognition is distributed over time, which includes short and medium-term actions to plan and prepare work, and long-term considerations such as how the system has evolved over time.

Each of these models has associated principles extracted from the literature on distributed cognition [10]. Table 1 shows how some of these are applied. These distributed cognition principles have been shown to help analysts gain further insight into complex sociotechnical systems compared to contextual design alone [14]. These principles should not be seen as a comprehensive and prescriptive check list of distributed cognition features, but as a set of sensitizing concepts to enrich what can be seen and described – some are likely to be relevant in a given context and some will not be. This approach has been applied in different contexts by our group and others (e.g. [15, 16]).

DiCoT-CL is an extension of DiCoT, which adds 'concentric layers' to each of the models to encourage the analyst to think about micro, meso and macro layers of distributed cognition [17]. DiCoT-CL has been applied to evaluate the design and use of a blood glucose meter on a ward [17] and to investigate the safety around infusion practices on a ward [18]. In both cases we treat the micro as the layer that is closest to the interactions to do with the device, procedure or technology under study, e.g. by the bedside; the meso is the layer out from this which might include different professionals as a team, e.g. at the scale of the ward; and macro is the layer that might be as broad as the hospital or above, e.g. national guidance. These different layers have been effective in showing that determinants for success and failure in a system might not be proximate to technology use but might be further away in space and time. For example, the configuration of infusion pump alarms in the hospital layer had downstream consequences for staff and patients in the micro layer [18].

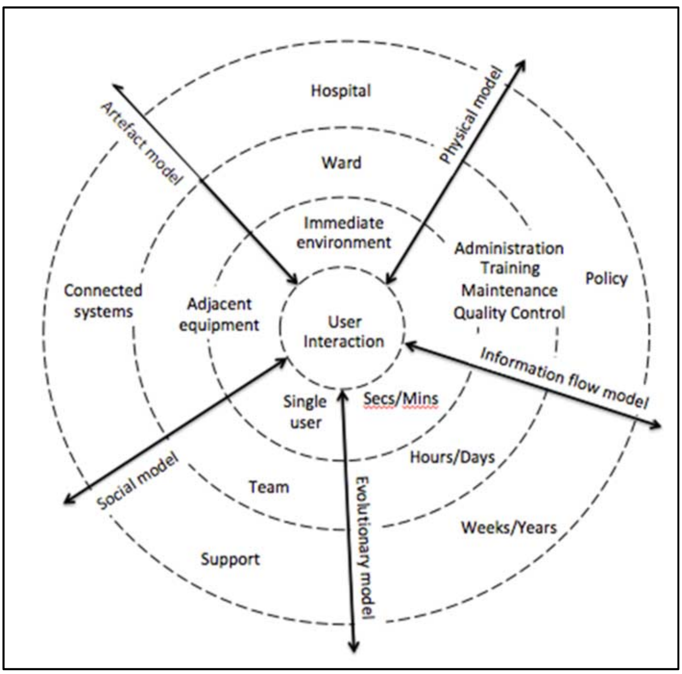


Figure 1. DiCoT-CL framework has three concentric layers of the sociotechnical system, where each layer is divided into five areas that reflect the themes of the different DiCoT models, i.e. information flow, artefact, physical, social and evolutionary models (adapted from [17]). Reproduced with permission.

Halverson [19] makes the point that distributed cognition does not explicitly name its concepts and constructs, which makes it harder to apply. DiCoT and DiCoT-CL advance the field by making explicit the concepts that support reasoning about a system in terms of distributed cognition, which are called 'principles' in DiCoT (e.g. see application of a sample of distributed cognition principles in Table 1). These principles are organised under the five DiCoT models that act as constructs for distributed cognition. DiCoT-CL adds constructs in the different layers of the sociotechnical system, and the proposition is that these layers are part of and nested within another. Naming these principles encourages deeper insight because they refer to more abstract sociotechnical issues, configurations and patterns in the data. The principles go beyond mere description and improve the rhetorical power of distributed cognition because the theory can help frame arguments about what is influencing the performance of the system, e.g. decisions in one layer might have an impact on performance in another layer.

Hollan et al. [2] offer a framework describing how integrated research activities using distributed cognition theory might be carried out. The framework links distributed cognition theory to ethnography, experimentation and testing, work materials, and workplaces. The order of these elements in the framework can be adapted for different studies; for example, observational work to identify key determinants for system performance might be conducted to inform the design of experimentation and testing; or it might be that results of tests or design processes away from the workplace can then be trialled to see how they perform in the messiness of practice.

2. Use of distributed cognition in health informatics

In this section we give a brief overview of how distributed cognition theory has been used in health informatics before presenting two use cases to demonstrate how DiCoT and DiCoT-CL can be used in both formative and summative stages of a project life cycle.

2.1. Overview of how distributed cognition theory has been used in health informatics

The case for distributed cognition's relevance for health informatics has been argued previously [20]. It lends itself well to exploring how people interact with technology in the workplace, issues to do with communication and coordination, how people's thinking extends into the environment and sociotechnical system architecture and performance more broadly. Published studies using distributed cognition in health informatics journals and related areas include:

- Hazlehurst et al. [21], who used distributed cognition to help identify six types of communication exchange that help situation awareness arise from coordinated work and achieve successful performance between surgeons and perfusionists in cardiac surgery. They investigated activities, artefacts, resources, constraints and information flows.
- Cohen et al. [22], who used distributed cognition to analyse morning rounds and handovers in a psychiatric emergency department to explore how error commission, detection and recovery are an integral part of cognitive work. They identified instances of perceived violations and miscommunication.

- Hussain and Weibel [15], who used DiCoT to explore the information flow of infection control information in critical care and consider the design of a novel touch screen and badge reader.
- Sarcevic and Ferraro [23], who examined the efficiency of electronic documentation in a fast-paced medical setting through the lens of distributed cognition, using cognitive ethnography to understand the information flow among team members and how information was shared, stored and documented.

2.2. Two use cases using distributed cognition for health informatics

We present two use cases at different stages of the project life cycle: the first case at the formative stage of the project life cycle where we are still developing an understanding of the key determinants for what, how, when and why patients use patient-held information about their medication (PHIMed); the second is an example of a study at the summative stage of a project life cycle where we evaluate the design and use of infusion devices already in practice.

2.2.1. Use case 1: Patient-held information about their medication (PHIMed)

PHIMed is defined as any patient-held information such that an editable list of current medications can be carried, regardless of whether or not other functionalities are also available; this can include paper and electronic tools [24]. Patients can be concerned about receiving appropriate treatment, especially where information breakdowns might occur in a fragmented healthcare system. For example, general practitioner surgeries, hospitals, and community pharmacies do not routinely share data in the National Health Service, increasing the chance of error. PHIMed helps fill the gaps and prevent errors. A specific example is that as a result of viewing a patient's PHIMed a community pharmacist stopped the patient purchasing an over the counter medication which would have been unsafe for them to use. [25] However, how PHIMed is best used, whether it affects patient outcomes and its key determinants for success are not known.

Distributed cognition was selected as the theory for investigating design and use of PHIMed. Distributed cognition seemed like a good fit since we are interested in the design and use of an artefact (i.e. PHIMed), how this may play a role in supporting cognition and decision making of healthcare professionals and patients, and how the use of PHIMed could have wider effects on the system such as improving resilience to errors and enhancing patient activation.

DiCoT is being used to provide concepts and constructs to inspire thoughts and questions that we might not have asked without the theory. Early on in the study we discussed the application of the theory and used it to brainstorm whether and how concepts and constructs might be applicable (see Table 1 for examples). These early ideas informed data gathering and analysis, but importantly will also be tested by empirical data. In this way DiCoT supported sense making at the very early stages of the project lifecycle, helping us to generate questions and explore potential determinants of success for PHIMed that can then be tested.

Table 1. Examples based on the PHIMed project showing how DiCoT principles can generate thoughts and questions before a project has started. The models show the area of Figure 1 that these principles are related to, however they are interdependent models so in practice there is a lot of overlap between models.

Principle (and associated model)	**Description of principle**	**Application to PHIMed project**
Information hubs (Information flow model)	Information hubs are a central focus where different information channels meet and different information sources are processed together, e.g. where decisions are made on various sources of information.	Patients may consider their general practitioner or healthcare record as a hub. However, others may see their care as fragmented so they are the hub, or they may perceive their PHIMed to act as a hub where there medications are concerned.
Behavioural trigger factors (Information flow model)	In teams it is possible for individuals to operate without an overall plan as each member only needs to know what to do in response to certain localfactors. Individuals may also base their behaviour on their perception of local circumstances. These initiating factors can be dubbed 'trigger factors' because of their property to trigger behaviour.	What was the reason for starting to use PHImed? What triggers someone to show PHIMed to a particular healthcare practitioner or to take it to a particular consultation? What triggers someone to update their PHIMed?
Situation Awareness (Physical layout model)	One of the key things in shared tasks is to keep people informed of what is going on, what has happened and what is planned. This can be influenced by how accessible the work of the team is. For example, in large control rooms the fact that an operator is in one area may lead to the correct inference of what they are doing, as that area pertains to certain activities.	This seems really important because different healthcare practitioners may lack situation awareness and PHIMed can help improve situation awareness in fragmented pockets of care.
Horizon of observation (Physical layout model)	The horizon of observation is what can be seen or heard by a person. This will differ for each person in an environment depending on their physical location, the activities they are close to, what they can see and hear, and the manner in which activities take place. The horizon of observation of a person refers to the scope of information input, whereas situation awareness is about the inferences that are made from this information.	Different healthcare practitioners will have very different horizons of observation, as will the patient and carer. Different technologies and PHIMed may influence the healthcare practitioner's horizon of observation.
Socially distributed properties of cognition (social model)	"The performance of cognitive tasks that exceed individual abilities is always shaped by a social organisation of distributed cognition. Doing without a social organisation of distributed cognition is not an option. The social organisation that is actually used may be appropriate to the task or not. It may produce desirable properties or pathologies. It may be well defined and stable or may change moment by moment; but there will be one wherever cognitive labour is distributed, and whatever one there is will play a role in determining the cognitive properties of the system that performs the task" [1, pp 262].	The cognitive tasks involved in looking after a patient, which could include treating different conditions, appear to be distributed between different specialists and individuals. How are these individuals distributed and coordinated? What impact does this have on patient care and decisions about medication optimisation?

2.2.2. Use case 2: Safety around infusion devices on a haematology ward

This study sought to evaluate safety around infusion devices on a haematology ward [18]. Using DiCoT-CL we explored determinant factors in the safety of infusions on the ward at the micro, meso and macro layers of the sociotechnical system. We present this as an example of how distributed cognition can be used in the late stages of the project lifecycle, post-implementation, to discover interactions that impact the performance of technology in practice. We spent 120 hours on the ward, shadowing and interviewing nurses, and observing infusion preparation and administration on the ward. DiCoT-CL was used for data gathering and evaluation; i.e., we attended to the five models, three different layers and the distributed cognition principles.

Using DiCoT provided leverage to investigate the complex sociotechnical informatics that comprises infusion practice on the ward. We found that infusion safety was influenced by artefacts, social networks, the flow of information, the physical layout of the ward and interactions over time. Issues with design of infusion pump alarms, medication storage, prescription information, and hospital bed management systems were noted. Safety is affected by the co-evolution of structure (e.g. the design of pump alarms), agency and workarounds (e.g. telling some patients how to silence their alarms), and deviations (e.g. patients silencing the wrong alarms) [18].

3. Explanation of success or failure of health information systems

This section describes determinant factors in the success or failure for the above mentioned use cases and how distributed cognition theory has supported the investigation and explanation of such factors.

3.1. Use case 1: Preliminary results in what triggers the use of PHIMed

DiCoT helped us start to think about the sociotechnical system that PHIMed may be embedded within. For example, early empirical feedback suggests that not all patients use PHIMed (i.e. they have different information requirements) but where it is used it can be recognised as an important artefact for supporting information flow around patient care. Before empirical work began we were already thinking about what contextual factors might be a ‘behavioural trigger’ to use of PHIMed (see application of DiCoT principles in Table 1). One conjecture was that patients who do not use PHIMed may perceive their general practitioner and/or medical record as an effective ‘information hub’ for their care (Figure 2.a), and that patients who do use PHIMed experience a more fragmented healthcare service that does not have an effective information hub. In this latter scenario the patient realises that they are the common feature in different consultations and act more like an ‘information hub’ in the absence of an effective hub on the healthcare side (Figure 2.b). Figure 2 shows two contrasting social networks inspired by the social model of DiCoT, which lead to different thoughts about what factors determine when PHIMed might be used and in what circumstances PHIMed might be most useful. These different patterns of distributed cognition are important for reflecting on who to target, when and in what circumstances to improve adoption and use in a future PHIMed intervention. We are continuing to test these and other conjectures in focus groups and interviews with healthcare professionals, patients and carers.

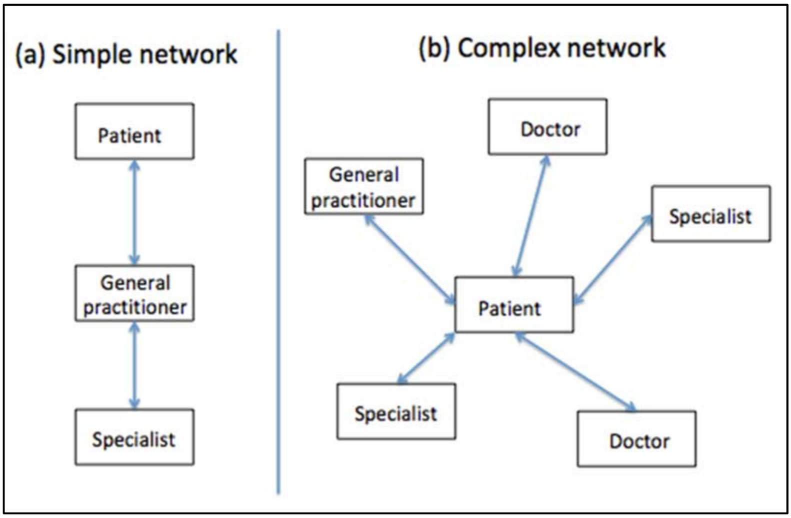


Figure 2. (a) An example of a simple social network where the general practitioner is seen as the central hub for patient information with only one other specialist; (b) An example of a complex social network where the patient is the central hub within fragmented healthcare services involving various doctors in secondary care and other specialists.

3.2. Use case 2: A problem between alarm design and barrier nursing

One of the most striking findings for safety around infusion devices on the haematology ward was a tension between the design of the infusion pump alarms and the procedures for barrier nursing [18]. The infusion pumps were designed to alarm ten minutes before the infusion was due to complete (a 'pre-alarm'). This was standard across the hospital and was intended to alert nurses so they could prepare for when it finished, e.g. prepare the next infusion. However, the haematology ward used many infusions and every patient needed barrier nursing to prevent spread of infection, requiring nurses to wash their hands and put on gloves, gown and mask before entering the patient's room. Reaching the pump to silence the alarm was therefore costly in terms of time. Once the nurse had silenced the pre-alarm they could not wait in the patient's room for ten minutes so they would go out and try to do something else before being called back for the actual alarm on completion of the infusion.

The physical layout of the ward also meant that it was difficult to hear pumps alarming from the corridor, partly because there was an ante-room between the main corridor and each patient's room. So, nurses were effectively relying on the pump alerting the patient, the patient pressing the call bell, and the nurse reacting to the call bell. Some patients expressed frustration at this process. One patient did not want to disturb the nurses, knowing how busy they are, so she sat next to the pre-alarm beeping for ten minutes before calling the nurse to attend to the pump.

A downstream consequence of the frustrations for staff and patients caused by the pre-alarm was that nurses would sometimes break with protocol and coach patients on how to silence the alarms. However, this would depend on the nurse and patient, e.g. patients who were not deemed competent would be discouraged from interacting with their infusion pump even if they tried to do it themselves. It was suggested by one member of staff that patients may have inadvertently silenced a low power warning thinking it was a pre-alarm, leading to the infusion pump stopping.

DiCoT-CL helped to describe the artefact, physical layout and social interactions that were contributing to this issue. There were also factors at play in different layers of the sociotechnical system. The main issues were being experienced at the micro and meso layer on the ward. Nurses complained that it was the way the pumps were designed, and there was nothing that could be done about it at the macro layer. However, further investigation revealed that this was a design configuration issue that could be adjusted by the hospital, at a different place in the macro layer. The analysis showed misunderstandings about the determinant factors contributing to this issue, which had different downstream consequences for staff and patients. Haematology services have now moved to a different hospital, but it seemed clear, through application of distributed cognition theory, that the configuration of the pre-alarms and barrier nursing was not working for staff and nurses. However, the pumps could have been reconfigured and tested using staff and patient satisfaction scores; efficiencies could have also been gained in the reduction of alarms. Again, distributed cognition can lead to testable hypotheses.

4. Discussion

Distributed cognition theory is relevant for the understanding of complex sociotechnical informatics. Cognitive concepts can be applied to sociotechnical systems to elucidate how information flows across artefacts, social networks, over different physical configurations and different spans of time. The theory had been around since the nineties, but more detailed support to facilitate its application has only been developed more recently. Many applications of distributed cognition theory provide description and explanation of how information flows in a sociotechnical system. However, it also lends itself to considering determinant factors for the success and failure of technology, and how the technology might fit (or not) in context. This can lead to design ideas and hypotheses that can be tested. For example, the use cases presented here show how distributed cognition can inform the design of PHIMed and in what circumstances patients might be most receptive to it; and how distributed cognition can be used to evaluate systems that have been already been deployed.

Like all theories, distributed cognition has strengths and weaknesses. Distributed cognition encourages a level of description about a system or process that lends itself to developing design ideas, but it may not readily emphasise the role of individuals or emotions as it focuses on systems and more observable functional issues [19]. This potential limitation is partially dependent on how rigidly one applies the theory. For example, more inductive ethnographic techniques can include interesting phenomena that might not readily fit the theory. Similarly, DiCoT can be applied in more or less rigid ways. Those unfamiliar with DiCoT may like to treat its models and principles as a check list, but we recommend its use in a semi-structured way to enrich what can be seen in context but not be too limited by the theory. Even with semi-structured use, DiCoT adds structure that draws attention to certain features and away from others, which should be reflected on by those who use it. Halverson [19] says that it is not clear whether success from the theory is from the theory itself or its commitment to ethnographically collected data. However, at least in the PHIMed example above we have seen that distributed cognition concepts and constructs can inspire thoughts and questions before empirical work has begun (Table 1), which suggest value lies in the theory itself. Since the theory is more accessible through the articulation of its concepts and constructs, and the need for such theories in health informatics and other domains is rising due to growing complexity of joint social and technical systems, distributed cognition is ready and relevant for grappling with twenty first century issues.

Following Halverson's [19] four categories for assessing the utility of theory: we believe distributed cognition provides *descriptive power* (e.g. DiCoT helped describe and make sense of activities through distributed cognition concepts and principles); *rhetorical power* (e.g. in terms of argumentation DiCoT-CL helped frame and articulate how macro level decisions were having negative downstream impact on the haematology ward); *inferential power* (e.g. we infer that patients who are most likely to use PHIMed would be those who experience complex and fragmented healthcare services); and *application power* (e.g. targeting PHIMed interventions towards patients with complex and fragmented care could have a higher likelihood of impact and changing the configuration of the pre-alarm on the haematology ward could have a positive impact on staff and patients, both of which could be tested).

Teaching questions for reflection

1. What tools and artefacts related to health informatics exemplify how an individual's cognition extends into the environment?
2. What complex sociotechnical systems can you describe that have important aspects of health informatics within them?
3. How would a research project evaluating electronic health records look different when planned from a traditional individualistic perspective versus a distributed cognition perspective?
4. What are the strengths and weaknesses of distributed cognition theory?

Funding

This work was supported by Pharmacy Research UK grant [PRUK-2016-PG2-2-A], by the UK Engineering and Physical Sciences Research Council [EP/G059063/1] and by the National Institute for Health Research (NIHR) grant [12/209/27], from the Health Services and Delivery Research (HS&DR) stream. The research is also supported by the NIHR Imperial Patient Safety Translational Research Centre. The views expressed are those of the authors and not necessarily those of Pharmacy Research UK, the NHS, the NIHR or the Department of Health and Care.

References

[1] E. Hutchins. *Cognition in the Wild*. MIT press.1995

[2] J. Hollan, E. Hutchins, D. Kirsh. Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction*, **7(2)**, (2000), 174-196.

[3] D. Furniss, A. Blandford. Understanding emergency medical dispatch in terms of distributed cognition: a case study. *Ergonomics* **49** (2006), 1174–203.

[4] N.V. Flor, E.L. Hutchins. Analysing Distributed Cognition in Software Teams: A Case Study of Team Programming during Adaptive Software Maintenance. Reading in Groupware and Computer supported Cooperative Work. San Mateo, CA: Morgan-Kaufman. 1991.

[5] E. Hutchins. How a cockpit remembers its speeds. *Cognitive science*, **19**(3), (1995), 265-288.
[6] D.K. Maglio, D. Kirsh. Some Epistemic Benefits of Action: Tetris, a Case Study. In Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society: July 29 to August 1, 1992, Cognitive Science Program, Indiana University, Bloomington 1992 (Vol. 14, p. 224). Psychology Press.
[7] J. Zhang, D.A. Norman. Representations in distributed cognitive tasks. *Cognitive science*. **18** (1994), 87-122.
[8] Y. Rogers. Recent theoretical developments in HCI: their value for informing system design (2000). Cited in Halverson C.A. Activity theory and distributed cognition: or what does CSCW need to DO with theories?. *Computer Supported Cooperative Work (CSCW)* **11** (2002), 243–67.
[9] P.C. Wright, R.E. Fields, M.D. Harrison. Analysing human–computer interaction as distributed cognition: the resources model. *HCI J* **15** (2000), 1–41.
[10] A. Blandford, D. Furniss. DiCoT: a methodology for applying distributed cognition to the design of teamworking systems. In: Interactive systems. Design, specification, and verification. Berlin Heidelberg: Springer; 2006. p. 26–38.
[11] J. Galliers, S. Wilson, J. Fone. A method for determining information flow breakdown in clinical systems. *Int J Med Informatics* **76** (2007), S113–21.
[12] N.A. Stanton. Representing distributed cognition in complex systems: how a submarine returns to periscope depth. *Ergonomics*. **57** (2014), 403-18.
[13] H. Beyer, K. Holtzblatt. Contextual design: defining customer centred systems. San Francisco, CA: Morgan Kaufmann Publishers; 1998.
[14] E. Berndt, D. Furniss, A. Blandford. Learning contextual inquiry and distributed cognition: a case study on technology use in anaesthesia. *Cogn Technol Work* (2014), 1–19.
[15] M. Hussain, N. Weibel. Can DiCoT Improve Infection Control?: A Distributed Cognition Study of Information Flow in Intensive Care. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (2016), 2126-2133.
[16] J. Rybing, H. Nilsson, C.O. Jonson, M. Bang. Studying distributed cognition of simulation-based team training with DiCoT. *Ergonomics* **59** (2016), 423-434.
[17] D. Furniss, P. Masci, P. Curzon, A. Mayer, A. Blandford. Exploring medical device design and use through layers of distributed cognition: how a glucometer is coupled with its context. *Journal of biomedical informatics*, **53** (2015), 330-341.
[18] D. Furniss, A. Mayer, B.D. Franklin, A. Blandford. The roles of structure, agency and deviations in safety: an ethnographic study of practices around infusion devices using distributed cognition and structuration theory. (submitted to *Safety Science*)
[19] C.A. Halverson. Activity theory and distributed cognition: or what does CSCW need to DO with theories? *Comput Supported Cooperative Work (CSCW)* **11** (2002), 243–67.
[20] B. Hazlehurst, P. Gorman, C. McMullen. Distributed cognition: an alternative model of cognition of medical informatics. *Int J Med Inform* **77** (2008), 226–34.
[21] B. Hazlehurst, C. McMullen, P. Gorman. Distributed cognition in the heart room: how situations awareness arises from coordinated communications during cardiac surgery. *J Biomed Inform* **40** (2007), 539–51.
[22] T. Cohen, B. Blatter, C. Almeida, V.L Patel. Reevaluating recovery: perceived violations and preemptive interventions on emergency psychiatry rounds. *J Am Med Inform Assoc* **14** (2007), 312–9.
[23] A. Sarcevic, N. Ferraro. On the use of electronic documentation systems in fast-paced, time-critical medical settings. *Interacting with Computers*. **29** (2017), 203-19.
[24] S. Garfield, D. Furniss, F. Husson, M. Turley, B.D. Franklin. Use of patient-held information about medication (PHIMed) to support medicines optimisation: protocol for a mixed-methods descriptive study. *BMJ Open*. **8** (2018), e021764.
[25] B. Jubraj, S. Barber, E. Ward, F. Husson, M. Turley. Patient perspectives on 'My Medication Passport'. [https://www.researchgate.net/publication/308765549_Patient_perspectives_on_'MY_medication_passport']

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

doi:10.3233/SHTI190114

Using Actor-Network Theory to Study Health Information Technology Interventions

Kathrin CRESSWELL[a,1]
[a] *Usher Institute of Population Health Sciences and Informatics, The University of Edinburgh, UK*

Abstract. This chapter introduces Actor-Network Theory, a sociotechnical approach to studying health information technology implementation. The chapter is intended as a pragmatic introduction to the field, acknowledging that there are many contested features of an Actor-Network Theory informed methodology. Nevertheless, the approach can be usefully drawn on to help to focus data collection and sampling. A case study describing the application of Actor-Network Theory to study the "failed" implementation of national electronic health records in England as part of a national "top-down" implementation program illustrates the main tenets of the approach and provides concrete examples of how Actor-Network Theory may be applied. In doing so, this chapter offers a reflexive account of how Actor-Network Theory has provided a nuanced analysis of how the implementation of national electronic health records affected different stakeholders, organizations and technology.

Keywords. Sociotechnical, Actor-Network Theory, Health Information Technology

Learning objectives

After reading this chapter the reader will be able to:

1. Describe the basics of Actor-Network Theory
2. Pragmatically apply Actor-Network Theory-based approaches to health informatics evaluations
3. Critically evaluate the various assumptions comprised within the Actor-Network Theory-based approach and draw on these for applied use in healthcare settings

1. Introduction to sociotechnical perspectives and Actor-Network Theory

1.1. Sociotechnical approaches to studying technology implementation

The concept of sociotechnical systems[2] emerged from the study of organizational behavior and workplace safety in 1950s studies of English coalmine workers.

The central assumption of sociotechnical approaches is that social and technical dimensions are intimately intertwined and need to be considered together when exploring

[1] Corresponding Author, Kathrin Cresswell, E-mail: Kathrin.Cresswell@ed.ac.uk
[2] See also Chapter 7, "Distributed Cognition: understanding complex sociotechnical informatics".

organizational dynamics and change. Humans in complex organizational contexts are viewed as being influenced by technological, social and cultural environments which are constantly changing. This is in turn assumed to affect the performance organization. In order to improve performance of the information system and the social context, it is argued that social and technical dimensions need to be aligned - a process referred to as joint optimization [1].

Sociotechnical perspectives have been applied to:

1) information system engineering - in order to inform technological designs that integrate well with the social environments in which they are used; and

2) the evaluation of information system implementation, adoption and optimization - in order to inform organizational efforts to create social environments that maximize benefits of technology [2].

This approach contrasts with earlier perceptions that workers had to adapt to technological requirements in order to realize potential benefits associated with technological change [3].

Sociotechnical approaches are a popular choice when examining the implementation of information technology (IT) in healthcare settings. They are well suited to explore changing organizational and healthcare professional practices accompanying technology introduction in complex environments [4-6]. Here, new technological systems are viewed as an addition to established organizational structures and work practices. These changes can lead to requirements for re-organization in social subsystems, often resulting in a tension between the technological demand for structured data entry and the fluid context-dependent nature of healthcare professional work. Conversely, sociotechnical approaches can also be used to examine how social and organizational practices result in changes in technological design.

Box 1 summarizes Coiera's four rules for sociotechnical design in healthcare settings, which exemplify the implications of a sociotechnical approach [2].

Box 1: Coiera's four rules for sociotechnical design

Rule 1: "Technical systems have social consequences": technology introduction affects the user and the individuals surrounding the user e.g. the patient

Rule 2: "Social systems have technical consequences": e.g. organizational culture, champions, role models may impact on the uptake of technology

Rule 3: "We don't design technology, we design sociotechnical systems": design needs to shift to incorporate social processes, and from consideration of a single user to the recognition that collaboration in healthcare is crucial

Rule 4: "To design sociotechnical systems, we must understand how people and technologies interact": the need to gather more data on human computer interaction in clinical environments e.g. cognitive overload, time pressured situations, workload

1.2. Actor-Network Theory-based approaches to studying technology implementation

Actor-Network Theory (ANT) can be viewed under the sociotechnical systems umbrella, as it focuses on exploring the interrelated nature of social and technological subsystems. It has its origins in sociological and anthropological approaches to organizational studies and was developed by science and technology scholars Bruno Latour, Michael Callon and John Law. ANT researchers view the world as consisting of networks made up of human and non-human actors. Non-human actors include objects and concepts. The central (and in some cases seen as somewhat controversial) assumption is that non-human and human actors should be treated as equal and that objects have agency (i.e. an ability to exert power over or change in other non-human or human actors) [7]. The ability to have power is assumed to emerge from the way actors are connected and is not assumed to stem from inherent actor characteristics [8]. However, this agency is not inherent to objects on their own – rather it emerges from the way they are related to other objects, concepts and human actors in the network. Both human and non-human actors can be a component of a network but also a network in themselves, depending on the level of granularity the ANT researcher wishes to study. Box 2 provides an overview of the most common terminology used in ANT. The terms provide an overview of the principles of ANT and they are listed for easy reference.

Box 2: Key terminology used in ANT

Actor: the origin of action (can be human or non-human)

Network: relationships between actors

Black-boxing: treating a particular network as a separate unit and specifying inputs and outputs as well as their relationship with the whole network

Intermediary: an individual that serves as a connection between two actors

Translation: process by which actors configure and re-configure each other

Simplification: composition of networks tends to reveal itself when things in a network go wrong and that they tend to be hidden when things go smoothly

Punctualization: process of revealing simplifications

In ANT, social phenomena are assumed to be the outcome of associations between actors [7], and the sociologist studying networks is simply seen as a component of the network. Therefore, ANT in its original "purist" form has been viewed as incompatible with interpretivist sociological approaches [7].

ANT scholars study the makeup and the shifting nature of networks and their components [3,9]. Typically, this involves focusing on some goal-directed collective activity, mapping network components, and in some cases specifying network inputs and outputs [3,8]. Based on this, it is assumed that researchers can make recommendations on how networks can achieve stability and how actors need to be re-configured to achieve a certain organizational aim. The stability of networks is assumed to be determined by the strength of relationships between actors [10].

Just as in overall sociotechnical approaches, networks are assumed to change and re-configure with the introduction of new technology (a new actor) in the organization [11]. Through tracing networks and investigating how they overlap and come into being, it is assumed that researchers can understand how power and organizational processes are generated [10,12].

1.3. The contested nature of Actor-Network Theory and its limitations

ANT is constantly evolving as it gets interpreted and re-interpreted by different scholars, which makes it somewhat hard to define its nature at any point in time [7,10]. It also has several limitations, and these have been extensively debated in the literature. The notion of non-human actors and their ability to possess agency is a particularly hot topic, with some doubting the contribution of this notion.[13-16] Whilst it is beyond the scope of this chapter to discuss all of ANTs shortcomings in detail, the most pertinent ones in relation to health technology implementation are discussed below.

Most importantly, it has been argued that ANT is not much of a theory at all as it lacks predictive power.[17,18] Predictive power is the ability of a theory to prospectively predict a phenomenon under investigation. In relation to health technology, this may for instance include postulations about how certain design features can lead to certain workarounds of users. According to Wacker, for a theory to be “good” it needs to have internal consistency and empirical riskiness - these are areas that ANT does not fare very well in.[18] Internal consistency refers to a theory providing logical and adequate explanations of reality. However, despite providing a vocabulary to describe social phenomena, ANT lacks the ability to explain and integrate the relationships between various human and non-human actors. As a result, ANT accounts can describe how clinical users and technology are related but may leave the reader questioning the actual contribution of applying the lens. Empirical riskiness encompasses the need for a “good” theory to be both risky and testable – but ANT cannot really be tested and lacks specificity.[18] Its terminology (see Box 2) is extremely loosely defined and its networks are potentially limitless, which can result in a lack of focus.[19,20] Consequently, if a theory cannot be subject to prospective tests, it may have limited usefulness.

Other criticisms have included the following: [10,21]

- ANT’s perceived inappropriate equal treatment of both human (e.g. clinical users) and non-human actors (e.g. technology, paper);
- the inability of the human observer/researcher to be fully agnostic (as postulated by ANT); and
- the lack of attention to the role of macro structures (e.g. economic, political environments) in influencing micro contexts (e.g. work practices, usability).

Nevertheless, ANT can help to facilitate interpretations of the researcher and provides a helpful vocabulary that can be used to explore a view of a world consisting of networks. This view of the world has several advantages when exploring the implementation of health IT. As such, ANT may be most appropriately viewed as a tool for theory development or a methodological approach in evaluating technology implementations.

2. Using Actor-Network Theory to evaluate health information technology

ANT has been employed by several medical sociologists to explore how artifacts and technologies can shape social processes in healthcare settings. In doing so, it has been applied rather pragmatically as a lens to examine specific aspects of technology implementation, to explore the effects of technological systems on human actors, and to explain why information systems may be rejected by users.

Although based on paper systems, perhaps the most illuminating examples of employing ANT can be found in a series of three case studies by Marc Berg and colleagues. These, drawing on both single physician-patient encounters and multi-disciplinary care teams, explore how the medical record actively impacts on human action and interaction.[4-6] Berg and colleagues provide detailed accounts of how the record constructs the patient's body/history and associated user practices, how it interconnects activities and actors through time and space, how it shapes relationships between actors and social processes, and how it serves different functions for different agendas. These agendas need to come together for the record to function. Berg at al. describe the complexity and situational ever-changing role of the record by focusing on detailed case scenarios. In doing so, they discuss connections between human actors and the record that capture the processes of how the two relate to each other in both formal and informal work practices.

Compared to paper records, electronic systems tend to pose greater challenges. They can connect human actors beyond physical space and can mediate a greater range of medical activity in a much shorter space of time. ANT has therefore also been used to explain why information system implementations in healthcare fail or why their adoption is often slower than anticipated. For example, Doolin and McLeod describe the implementation of an executive information system in a hospital in New Zealand.[10] The authors argue that failure to build the new network (i.e. the information system) was due to "an inability to enroll the non-human actors" (p.259), which in this case consisted of a perceived lack of data quality in the new digital system. Hence, its use was rejected by doctors.

Similarly, Whitley and Pouloudi use ANT as a framework for analyzing the introduction of NHSnet.[22] NHSnet is a Microsoft Outlook based Web App system that supports communications of medical information in the United Kingdom (UK) National Health Service (NHS). Implementation was completed on time but there were heated discussions with the medical profession over confidentiality and security issues surrounding medical information. In this context, ANT helped the authors to conceptualize how different human stakeholder groups (including doctors, professional groups and technologists) have different interests that are not easily aligned within a single technological solution. As a result of ongoing discussions, NHSnet's design changed over time. This in turn had implications for how human actors were positioned in the network.

2.1. Drawing in Actor-Network Theory to explore the national implementation of electronic health records in English hospitals

In our own work, we have used ANT to examine the implementation of electronic health records (EHRs) in hospitals as part of the English National Programme for Information Technology (NPfIT) (see Figure 1 and Box 3). This case study will be used to illustrate how ANT can helpfully be applied to inform data collection and analysis in studies of health IT implementation.

Box 3: Summary of an evaluation of hospital EHR introduction in the NPfIT [23]

- National implementation of centrally procured software in hospitals
- Qualitative longitudinal investigation in three purposefully selected hospitals which were implementing early functionality (conceptualized as case studies)
- Collected data between 2009 and 2011
- Dataset: 66 interviews with hospital staff, 14 interviews with stakeholders from outside case study sites, 38.5 hours of non-participant observation, 149 pages of press statements, 31 pages of field notes, and a range of national and local documents
- Key findings: users found it difficult to integrate the software with their everyday work practices as the software was perceived to be not fit-for-purpose, implementation had significant consequences for organizational functioning (hampered by local restrictions in software customizability)

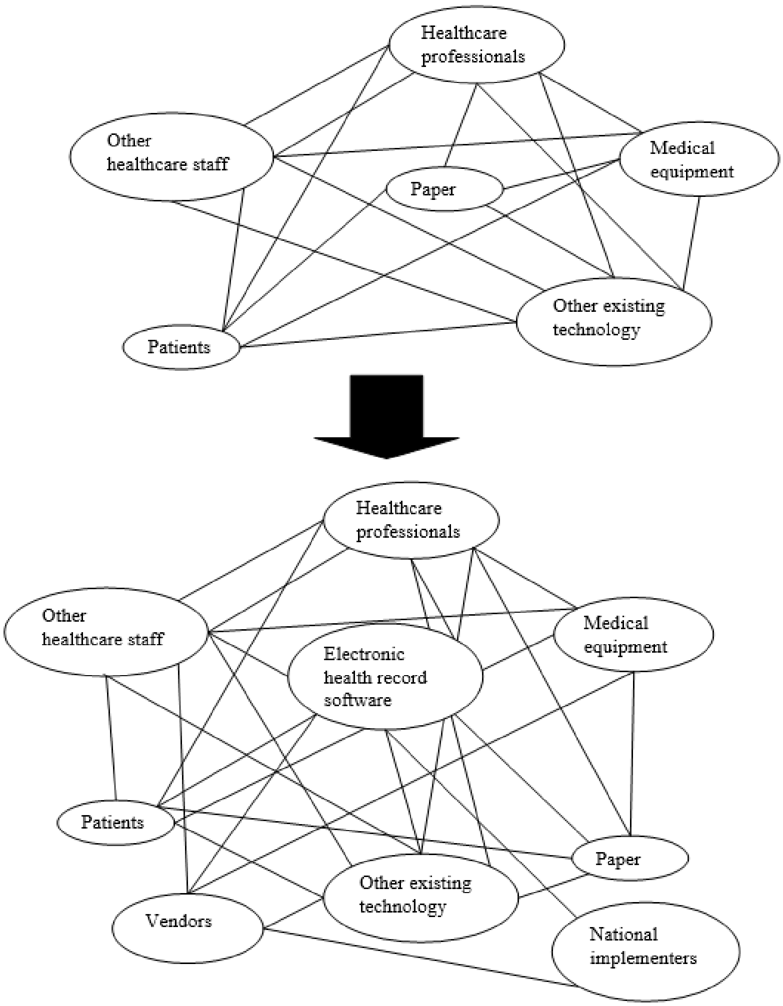


Figure 1. Illustration of an Actor-Network in evaluating the introduction of EHRs in English hospitals as part of the National Programme for Information Technology (based on and partly reproduced from [19].

The context for this analysis was the national procurement of EHRs in English hospitals by the government in 2002. Three commercial information system suppliers were tasked with delivering these, driven by a vision to deliver a record that connected disparate sources of information across care settings on a national scale. In 2011, the £12.7 billion NPfIT initiative was abandoned.

A pragmatic ANT-informed approach was used to explore how EHRs transformed care, healthcare staff relationships, and wider macro-networks (including policy makers and supplier relationships). Despite its limitations outlined above, the notion of the active role of objects in shaping human relationships was a helpful lens to examine human dynamics and technological changes over time. EHRs, other technology, and paper records were viewed as non-human actors, whilst attempts were made to black-box the hospital EHR and analyze its translations over time (see Figure 1).

ANT was used as a conceptual tool for focusing data collection efforts. This was done through sampling those human actors that were connected to the EHR and tracing their relationships (see Figure 1). These network connections were either direct by interacting with or building the technology, or indirect by influencing its strategic direction. Over time, it became clear that the network was not confined to the hospital environment but included for instance policy makers that had procured national technological systems, the media, and information system developers. These could be viewed as intermediaries, as they had an indirect influence on how the technology was used by healthcare professionals.

ANT also helped to conceptualize how care was organized around the record and how the re-organization of the record (by making it electronic) in turn re-organized care and healthcare professional roles (translation). Due to the electronic nature of the EHR, this change was often done at a distance thereby connecting spatially disconnected areas of care. Ultimately, the vision was to do this nationally. Such relational connections and mechanisms are usually poorly described and mapping network components helped to reveal these. Investigating these processes, however, can have practical implications and helped to identify facilitators for adoption and implementation that may otherwise not have been considered. The most helpful aspect here was that ANT facilitated viewing the EHR as an active part of the social world. As in Berg's examples outlined above, we examined how the technology influenced the social relationships of healthcare staff using it, information system developers, patients, policy makers and evaluators (see Figure 1). We also explored how the introduction of technology resulted in the formation of new networks and how these transformed over time (translation). Here, ANT helped to conceptualize how individual and organizational practices were centered around the record, and how its role of directing organizational activity changed when it became electronic.

3. Explaining the "failure" of the nationally procured EHR drawing on ANT

ANT is not a theory in the traditional sense. It describes rather than explains and its explanatory power is limited. Nevertheless, it provided a unique and in-depth insight into the processes and the active role of the EHR in coordinating care and work practices of healthcare staff and hospitals. Thus, it helped to draw a sophisticated picture surrounding the implementation and adoption of nationally procured EHRs that went beyond the simple dichotomy of "success" and "failure". This is because new network formations can be described without making value judgements. Accordingly, analytical focus shifted from dichotomizing towards stakeholder sense making activities, to negotiation and aligning differing actor perspectives/behaviors.

A range of perspectives reflected different views surrounding "success" and "failure" resulting from different positions within the network at different points in time. For instance, the new software resulted in increased workloads for nurses, who may have viewed the implementation as a "failure" at least in the short- to medium-term. Policy makers, in turn, focused on the progress of developing the infrastructural components underlying the EHR technology. They therefore viewed this aspect of the national implementation as a "success".

The introduction of the new EHR also affected stakeholders in many different ways and revealing these simplifications was a key analytical task. The level of influence depended on their role (healthcare professionals, managers, policy makers, information system suppliers, patients), their local setting (existing relationships, physical environments), and the technology adoption time (short-, medium-, long-term). Common to all contexts and individuals, however, was that the technology adopted was an immature solution that lacked usability and had mostly negative effects. It for instance, increased workloads of users and negatively affected reputations of managers and suppliers. Over time, as the solution matured, some networks began to stabilize with the record gradually fulfilling its purpose of coordinating care effectively and stakeholders acclimatizing to these changes. However, these changes were only visible on a very small scale and in settings that had invested a significant amount of time and resources.

The new information system was procured nationally, so policy makers and system vendors were initially in a relative position of power. Over time, clinical users became more powerful, as they refused to use a technology that was viewed to lack usability. This changed power dynamic led to changes in the national procurement model. It is not to say that other sociotechnical dimensions (including other social, political and organizational factors) are not important in determining "success", but this work has illustrated that the most important pre-requisite for "success" from all perspectives is a usable technology.

When mapping out the larger network and tracing the technology, we ensured that all human actors were either directly or indirectly (i.e. through another actor) related to the EHR. ANT-informed analysis indicated that there were two different networks that were not effectively connected through strong associations (e.g. aligned interests) beyond the technology itself (see Figure 1). These were the users of the technology (i.e. healthcare professionals and organizations) and the national implementers (i.e. policy makers and information system suppliers). Both groups had different views of and intentions for the technology: Policy makers wanted to make or save money/lives on a large scale through improving organizational processes. Users wanted to improve immediate patient care in their own micro-environments. There was thus a tension between the micro and the macro networks in the following ways:

1. Policy makers and suppliers foregrounded the vision of the technology as an integrated national EHR, whilst users had to cope with its manifestations and its lack of usability in everyday life.
2. New technology was designed to structure care to make it more effective (including the imposing of rules, categories and regulations). This was at odds with the nature of clinical reality as these rules inhibited the timely provision of care and also increased individual workloads.

The lack of alignment of these positions was apparent in ongoing tensions and eventually broke up the network, reflected in the media discourse surrounding "the spectacular failure of the NPfIT".

4. Discussion

This chapter has illustrated that drawing on ANT can be helpful in conceptualizing technology implementation in healthcare settings. In particular, the approach can help to inform sampling and to examine how technology actively shapes human relationships and vice versa. It can further inform deliberations on the alignment of various networks at different levels including healthcare professional work, organizational practices, political and supplier relationships. In line with this, ANT-informed approaches continue to be used by health service researchers as tools to facilitate tracing networks of human actors and technologies over time. However, these are mostly small-scale studies exploring health IT implementations in particular settings [24-26].

Due to its limitations, the traditional "purist" ANT approach is likely to be too restrictive and too prone to getting lost in detail to be usefully employed in studying health IT implementation and this is reflected in the current literature, where the use of ANT to inform analysis is generally less common than its role in informing design considerations [24-26]. It is therefore often employed in conjunction with other theoretical lenses under the more general sociotechnical umbrella [20,25].

Sociotechnical lenses are well suited to explore processes surrounding technology implementation across a variety of different stakeholder levels [27,28]. These approaches are proving to be relatively flexible, in particular when considering large-scale implementations of complex programmes where drawing on one single lens may be quite restrictive [29]. However, many existing approaches still examine health IT at one selected level, be it micro contexts, meso (organizational) contexts or macro contexts [30-33]. The relationships between these are often poorly understood and this is where pragmatic ANT-informed approaches, as outlined in the case study above, may be useful for evaluators going forward.

It is difficult to predict if the NPfIT would have had more successes if the design of technologies and implementation strategy had drawn on ANT. This is because the application of the method is very much subjective. However, more generally, rigorous independent formative evaluation methods (informed by ANT in combination with other sociotechnical lenses) are crucial to accelerate learning and optimization of implemented technologies and practices.

Teaching questions for reflection

1. Consider which objects in your environment may be viewed as having agency.
2. Would you draw on ANT in your work? How could you do this?
3. What do you consider to be the most helpful/unhelpful aspects of ANT in health IT implementation?

References

[1] V.L. Mitchell VL, B.R. Nault, *The emergence of functional knowledge in sociotechnical systems. University of Calgary*, Haskayne School of Business, 2003.

[2] E. Coiera, Four rules for the reinvention of health care, *BMJ: British Medical Journal* **328** (2004), 1197.

[3] W.N. Kaghan, G.C. Bowker, Out of machine age?: complexity, sociotechnical systems and actor network theory, *Journal of Engineering and Technology Management* **18** (2001), 253-69.

[4] M. Berg, Patient care information systems and health care work: a sociotechnical approach, *International Journal of Medical Informatics* **55** (1999), 87-101.

[5] M. Berg, J. Aarts, J. van der Lei, ICT in health care: sociotechnical approaches, *Methods Archive* **42** (2003), 297-301.

[6] M. Berg, C. Langenberg, I. vd Berg, J. Kwakkernaat, Considerations for sociotechnical design: experiences with an electronic patient record in a clinical context, *International Journal of Medical Informatics* **52** (1998), 243-51.

[7] B. Latour, *Reassembling the social: An introduction to actor-network-theory*, Oxford University Press, 2005.

[8] M. Müller, Assemblages and actor ‐ networks: Rethinking socio ‐ material power, politics and space, *Geography Compass* **9** (2015), 27-41.

[9] A. Prout, Actor‐ network theory, technology and medical sociology: an illustrative analysis of the metered dose inhaler, *Sociology of Health & Illness* **18** (1996), 198-219.

[10] B. Doolin, L. McLeod. Towards critical interpretivism in IS research, *Handbook of Critical Information Systems Research* (2005), 244.

[11] B. Doolin, A. Lowe, To reveal is to critique: actor–network theory and critical information systems research, *Journal of Information Technology* **17** (2002), 69-78.

[12] J. Law, Notes on the theory of the actor-network: Ordering, strategy, and heterogeneity, *Systems Practice* **5** (1992), 379-93.

[13] E. Sayes, Actor–Network Theory and methodology: Just what does it mean to say that nonhumans have agency?, *Social Studies of Science* **44** (2014), 134-49.

[14] L. Baron, R. Gomez, The associations between technologies and societies: the utility of actor-network theory, *Science, Technology and Society* **21** (2016), 129-48.

[15] A. Whittle, A. Spicer, Is actor network theory critique?, *Organization Studies* **29** (2008), 611-29.

[16] G. Walsham, Actor-network theory and IS research: current status and future prospects, In *Information Systems and Qualitative Research* 1997 (pp. 466-480), Springer, Boston, MA.

[17] M. Callon. Actor-Network Theory – the Market Test, Department of Sociology, Lancaster University, available from: http://www.comp.lancs.ac.uk/sociology/papers/Callon-Market-Test.pdf (last accessed 09/04/2018).

[18] J.G. Wacker, A definition of theory: research guidelines for different theory-building research methods in operations management, *Journal of Operations Management* **16** (1998), 361-85.

[19] K.M. Cresswell, A. Worth, A. Sheikh, Actor-Network Theory and its role in understanding the implementation of information technology developments in healthcare, *BMC Medical Informatics and Decision Making* **10** (2010), 67.

[20] K. Cresswell, A. Worth, A. Sheikh, Implementing and adopting electronic health record systems: How actor-network theory can support evaluation, *Clinical Governance: An International Journal* **16** (2011), 320-36.

[21] D. Howcroft, E.M. Trauth, *Handbook of critical information systems research: Theory and application*, Edward Elgar Publishing, 2005.

[22] E.A. Whitley, A. Pouloudi, Studying the translations of NHSnet, *Journal of Organizational and End User Computing* **13** (2001), 30.

[23] A. Sheikh, T. Cornford, N. Barber, A. Avery, A. Takian, V. Lichtner, D. Petrakaki, S. Crowe, K. Marsden, A. Robertson, Z. Morrison, K. Cresswell, Implementation and adoption of nationwide electronic health records in secondary care in England: final qualitative results from prospective national evaluation in "early adopter" hospitals, *BMJ* **343** (2011), d6054.

[24] R.G. Booth, M.A. Andrusyszyn, C. Iwasiw, L. Donelle, D. Compeau D, Actor ‐ Network Theory as a sociotechnical lens to explore the relationship of nurses and technology in practice: methodological considerations for nursing research, *Nursing Inquiry* **23** (2016), 109-20.

[25] A. McDougall, M. Goldszmidt, E.A. Kinsella, S. Smith, L. Lingard, Collaboration and entanglement: An actor-network theory analysis of team-based intraprofessional care for patients with advanced heart failure, *Social Science & Medicine* **164** (2016), 108-17.

[26] L. Nguyen, L. Bakewell, N. Wickramasinghe, P. Haddad, I. Muhammad, H. Moghimi, D. Wilde, B. Redley, J. Considine, M. Botti, Transition from paper to electronic nursing documentation in residential aged care: An actor network theory analysis, *Electronic Journal of Health Informatics* **8** (2015), 4.
[27] A. Takian, A. Sheikh, N. Barber, We are bitter, but we are better off: case study of the implementation of an electronic health record system into a mental health hospital in England, *BMC Health Services Research* **12** (2012), 484.
[28] K. Cresswell, D. Bates, R. Williams, Z. Morrison, A. Slee, J. Coleman, A. Robertson, A. Sheikh, Evaluation of medium-term consequences of implementing commercial computerized physician order entry and clinical decision support prescribing systems in two 'early adopter' hospitals, *Journal of the American Medical Informatics Association* **21** (2014), e194-202.
[29] T. Greenhalgh, J. Wherton , C. Papoutsi, J. Lynch, G. Hughes, C. A'Court, S. Hinder, N. Fahy, R. Procter, S. Shaw, Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies, *Journal of Medical Internet Research* **19** (2017), 11.
[30] E. Kaldoudi, A. Chatzopoulou, V. Vargemezis, Adapting the STARE-HI guidelines for the evaluation of home care telehealth applications: An interpretive approach. *The Journal on Information Technology in Healthcare* **7** (2009), 293-303.
[31] P. Nykänen, J. Brender, J. Talmon, N. de Keizer, M. Rigby, M.C. Beuscart-Zephir, E. Ammenwerth, Guideline for good evaluation practice in health informatics (GEP-HI), *International Journal of Medical Informatics* **80** (2011), 815-827.
[32] J. Talmon, E. Ammenwerth, J. Brender, N. De Keizer, P. Nykänen, M. Rigby, STARE-HI—Statement on reporting of evaluation studies in Health Informatics, *International Journal of Medical Informatics* 78 (2009), 1-9.
[33] J. Warren, M. Pollock, S. White, K. Day, Y. Gu, A Framework for Health IT Evaluation, In *Health Informatics New Zealand 10th Annual Conference and Exhibition* 2011.

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

doi:10.3233/SHTI190115

Collective Mindfulness and Processes of Sensemaking in Health IT Implementation

Valentina LICHTNER[a b,1] and Johanna I WESTBROOK[b]
[a] *Department of Practice and Policy, UCL School of Pharmacy, UK*
[b] *Centre for Health Systems and Safety Research, AIHI, Macquarie University, Australia*

Abstract. High reliability organisations operate safely in situations of high risk by organising for collective mindfulness. They do so through five ongoing processes geared towards anticipating, containing, and making sense of the unexpected. The five processes are: preoccupation with failure, reluctance to simplify interpretations, sensitivity to operations, commitment to resilience, and deference to expertise. The theory of collective mindfulness builds on Hutchins's theory of distributed cognition (the 'collective mind' of ship navigation teams) and on Langer's theory of mindfulness about individuals' interpreting information in context. However, in the theory of collective mindfulness, attention is paid not to individual cognition or decision making, but to collective processes of sensemaking emerging from individuals' interactions in dealing with an equivocal environment. In health informatics, the theory of collective mindfulness can be used to explain health information technology (IT) development and implementation, across its life cycle, and inform guidance towards mindful management of IT projects. For example, applied to a case of electronic health record implementation in a hospital context, the theory explains how mindful management of the sense-making challenges of post-roll out adaptation processes contributes to a 'successful' IT project. Further, the theory challenges a static and linear understanding of success (or failure) of health IT initiatives, supporting instead an argument for outcomes – be it reliability and safety, or IT project success – as collective, complex and dynamic achievements of mindful organising practices.

Keywords. Mindfulness, Sensemaking, Organisations, Technology adoption.

Learning objectives

After reading this chapter, the reader will be able to:

1. Describe and explain the main tenets of the theory of collective mindfulness in organisations.
2. Translate and apply the theory to health information technology (HIT) implementation contexts.
3. Evaluate the strengths and weaknesses of the theory with respect to the insight it can provide on HIT implementation.

[1] Corresponding Author: Valentina Lichtner, E-mail: valentina.lichtner@mq.edu.au

1. Introduction: the theory of collective mindfulness

> *Mindfulness is a state of being attentive to new information, new meanings, and different points of view [1]. It is not about meditation.*

The theory of collective mindfulness originated in the mid-1990s from research that applied an organisational lens to the investigation of high reliability organisations (HRO) [2-4]. Among the originators of the theory were Karl Weick, Karlene Roberts, Kathleen Sutcliffe and other members of 'the HRO Project' at Berkeley (University of California). HROs are organisations operating in high levels of complexity and risk, but where serious accidents are extremely rare. Examples of these type of HRO are naval and armed forces, fire services. These organisations cannot afford to learn from trial and error as other organisations might do; instead, their 'first error is the last trial' [3][p32]. The theory of collective mindfulness arose from the investigation of these HROs as a way to explain how they are able to work in highly complex environments and yet ensure few major errors occur.

Collective mindfulness is an organisational state of being, or way of working, which is characterized by *'a quality of organizational attention that increases the likelihood that people will notice unique details or situations and act upon them'* [2][p410]. It emerges from five ongoing reliability-enhancing (collective) cognitive processes geared towards anticipating and containing the unexpected (Table 1): *preoccupation with failure, reluctance to simplify interpretations, sensitivity to operations, commitment to resilience,* and *deference to expertise* [2, 3]. The first three processes sustain an organisation's capacity to anticipate, and make sense of, 'the unexpected'. For example, to anticipate a fault in equipment. The last two processes focus on dealing with, and containing, the problem, before it results in an accident or harm. Through the enactment of these processes, an organisation shows the capacity for resilience.

Table 1. Five dimensions of *collective mindfulness* (adapted from [2, 3, 5])

Dimensions	Definition
Preoccupation with failure	An ongoing wariness that errors are possible; paying attention to things going right, those that do not go right, and how things could go wrong; small failures and near misses are treated as indicators of potentially bigger issues.
Reluctance to simplify interpretations	Not taking the past as the only guide to the present. Making fewer assumptions, questioning usual wisdom, uncovering blind spots, bringing more perspectives to achieve understanding.
Sensitivity to operations	Creating and maintaining an integrated 'big picture' of the current situation in the moment, through real time information. Similar to situation awareness, it involves the envisioning of possible future states and knowledge of interconnections.
Commitment to resilience	Awareness that it is impossible to eliminate uncertainty or anticipate all situations. Capacity building. The enlarging of individual and organisational capabilities to enable recovering from the unexpected (what cannot be anticipated). Capabilities include widening of 'repertoires of actions', skills at improvisation, 'recombination' and adaptation, ad hoc networks. Ways to achieve this include incorporating lessons from the past, training and learning from feedback.
Deference to expertise (also referred to as flexible decision structures or under-specification of structures)	Enabling the persons with the greater expertise to handle the problem and make decisions regardless of rank or hierarchy. This requires flexibility in organisational structures.

These processes are interrelated as part of a dynamic whole (Figure 1). Each of the five processes depend on the other and is maintained through feedback and learning. HROs 'socialize people to notice more' [3] over a background of constant preoccupation with 'things going wrong'; their organisational structures are compatible with maintaining and enhancing resilience. Once something is noticed, it is shared. Collective mindfulness depends on ongoing sharing of information, communication and interaction between individuals, so that interpretation of what is happening can be refined, beyond usual assumptions, and with awareness of overall workflows and interdependencies. This is thus a process of collective sensemaking – making sense of the situation overlaps with actions to solve or contain the problem, involving people with the right expertise, beyond hierarchical lines. Through the process, the organisation learns, broadens individual and organisational repertoires of actions, and gains collective knowledge that will inform the making sense of future uncertain, unexpected situations.

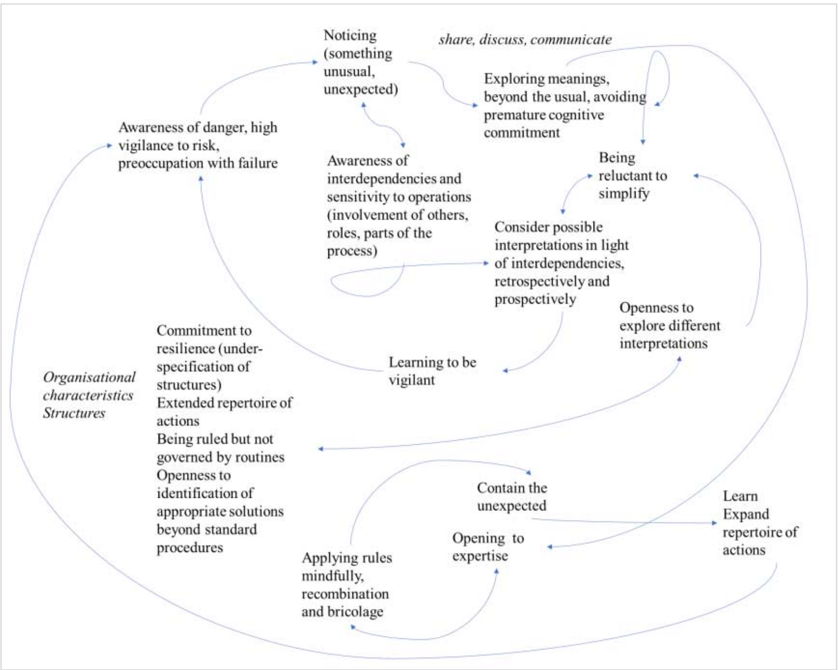


Figure 1. The collective mindfulness whole: a dynamic achievement

1.1. The cognitive dimension

The theory of collective mindfulness builds on the view of a 'collective mind' emerging from distributed processes proposed in distributed cognition [6]. Collective mindfulness also draws on Langer's theory of mindfulness about individuals' interpreting information beyond premature cognitive commitments [1]. However, collective mindfulness shifts the focus away from individual cognition, to collective processes of sensemaking emerging from interactions between people working in equivocal environments [7, 8]. Within this perspective, the 'collective mind' is 'embodied in the interrelating of social activities' [2, 4] and an organisation (or the group or team) can be said to be mindful, or have a mindfulness capability, the same way that it is said that organisations learn or have learning capabilities [9].

In contrast to other studies of cognition in organisations, understanding activities and organisations in terms of sensemaking and collective mindfulness takes us away from the traditional decision-making lens. Traditionally decision making is explained as a rational selection between options. Although seen as the combination of two processes of judgement and choice, decision making is typically studied from an individual (cognitive) perspective. The focus is often placed on the outcome (the judgment, the decision). Instead, understanding sensemaking requires a dynamic perspective and a focus on the (social) activity in context.

1.2. Making sense and 'the unexpected'

Sensemaking in organisations is about making sense of unusual events, ambiguous information, or unexplained variations of performance. Organisations are complex sociotechnical environments – they are open adaptive systems. Take for example the implementation of a large clinical information system being rolled-out across multiple hospitals. In such environments it is inevitable that new events and situations will appear that could not have been anticipated, or were known to some but not made known to the designers/implementers (also referred to as 'unknown knowns' [10]). These situations create uncertainty and therefore require 'making sense of'. This is what is usually referred to as *'the unexpected'* [8], although the expression may also be used more generally to refer to any unwanted outcomes or issues (errors, accidents) in organisational processes. Examples of the unexpected in an IT implementation project may include agreed requirements that become contested, unanticipated changes in workflows, or use of an IT system to complete work in ways which had not been foreseen. The case below illustrates making sense of, and resolving the unexpected in an IT initiative requiring the involvement of stakeholders, and processes of discovery and negotiation – overall, a process of collective sensemaking.

In the organisational literature, depending on the perspective taken to understand the making of sense in organisations, attention has been paid to activities of *sense-giving* (attempts to influence others' interpretations), *sense-breaking* (when sense is 'breaking down'), or *sense-exchanging* (social negotiation), among others. We will return to some of these in the case discussed below.

1.3. Methods used to research phenomena within this theoretical frame

Weick and Roberts explain that the word 'collective' *"refers to individuals who act as if they are a group".* This means they "interrelate their actions" (and they do so "with more or less care") [4][p360]. Since its beginning, empirical research on collective mindfulness has therefore attempted to capture these processes of interrelating, achieving this with in-depth ethnographic case studies (e.g. [4]). This method is the most suitable for capturing the dynamics of activities in context. In these studies, the unit of analysis often shifts between individuals and groups [3], *'since only individuals can contribute to a collective mind, but a collective mind is distinct from an individual mind because it inheres in the pattern of interrelated activities among many people'* [4][p360].

Qualitative studies of this kind have made useful contributions to the original theory by investigating these phenomena in very different organisational contexts. For example 'decision surprises' in the banking sector [11], and the dialectics of collective minding in a building project of a renowned architecture firm [12].

A different type of research on collective mindfulness aims to more formally assess and measure the association between antecedents (e.g. organisational structures) and mindful organising, and between mindful organising and organisational outcomes (reliability, safety) (Table 2). Vogus and Sutcliffe designed and tested a scale to measure levels of mindful organising across healthcare organisations (nursing units). This scale was then used to examine associations between levels of mindful organising with organisational outcomes (medication errors) [13]. More recently, the scale was used in a similar, more complex, mixed-method study where mindful organising was taken as one of four reliability-enhancing work practices (REWP) (in addition to: respectful interaction; affective commitment; and organisational citizenship behaviour). A survey of 10 items capturing information about each of these REWP was administered to nurses in 95 units across 10 hospitals. They were then correlated to outcomes of patient care (medication errors and patient falls). In assessing the constructs, the authors found mindful organising correlated with respectful interaction. However, mindful organising had significant correlation with medication errors and patient falls, while respectful interaction did not [14][p.14]. The study found that scores on 'mindful organising' were significantly negatively associated with medication errors and falls, i.e. units with higher mindful organising scores had significantly lower rates of medication errors and falls. [14][p16]. Thus, measurements of this kind add predictive value to the theory, i.e. making it usable to predict the likelihood of positive or negative outcomes in the presence of different levels of mindful organising practices. Their drawback is loss of insight into the dynamics of collective mindfulness processes that may be specific to different organisational contexts.

Overall, it is important to remember that studying, or measuring, collective mindfulness is not the same as studying, or measuring, individual mindfulness. The latter is increasingly of interest among researchers of users' behaviour with technology, and can be achieved, for example, by use of the generic Langer's mindfulness scales, or scales designed more specifically to capture individual's mindfulness with IT. But studying individual mindfulness does not reveal collective capabilities, organisational processes or outcomes (and there is no evidence yet that individual mindfulness directly produces collective mindfulness [2]).

Table 2. Examples of methods used to empirically study *collective mindfulness*

Aims	Examples of methods
Exploring how collective mindfulness unfolds and why; *understanding* the dynamic processes (and possible mechanisms) that lead to positive or negative outcomes (reliability/failures)	Qualitative research; in depth case studies.
Measuring collective mindfulness (presence/absence or levels)	Design of scales distributed and tested through surveys
Assessing/Testing association between organisational characteristics (antecedents) and mindful organising, or between mindful organising* and organisational outcomes (e.g. reliability/failures)	Scales and surveys; mixed methods.

* Mindful organising here is equivalent to 'an intervention' (e.g. to improve reliability or safety). However, to our knowledge there have been no studies that introduce organisational mindfulness as an intervention and then test its effects on outcomes.

1.4. Collective mindfulness and IT implementation

The theory of collective mindfulness has been applied in the field of information systems (IS) to both explain aspects of IT implementation and as a recommendation for improvement [15]. For example, in relation to information system development, Butler and Gray [9] suggest that a collective mindfulness approach may lead to more successful IT projects (better able to manage project risk). They argue that agile development techniques, in contrast to formal development methods, may promote mindfulness by focusing attention on 'what is needed and what exists, rather than the abstractions of what is expected or promised' (with risks of premature cognitive commitments) [9][p220]. Once systems are implemented and put into use, they are often found to be 'fundamentally unreliable' [9][p217]. Butler and Gray [9] argue that collective mindfulness can explain how organisations using such systems achieve reliability, for example, by mindfully managing business continuity and disaster recovery and organising operations of a 'technical support' unit. The authors also point out how paradoxically, systems designed for ease of use may have negative implications on users' ability to achieve reliable outcomes when they *'provide results tailored to one perspective, and avoid revealing alternative perspectives'* [9][p220]. Here they argue that such design approaches may hinder *collective mindfulness* when they 'mask unexpected variation', 'promote efficient routinized behaviour' and restrict choice [9][p221].

2. Use of collective mindfulness in health informatics

A rare case of the application of the theory of collective mindfulness in health informatics, is Aanestad and Jensen's study [16] of a Norwegian hospital adoption of an electronic health record (EHR) system. Their interest lies in the post-implementation adaptation processes, and in particular those changes that 'technology triggers' after implementation is officially over and 'the dust has settled' [16][p15]. With respect to the traditional life-cycle of an IT implementation, understanding these processes is an important part of the evaluation phase, where evaluation overlaps with, and informs, further adaptive design and development.

2.1. The case, as recounted in Aanestad and Jensen, 2016 [16]:

A Norwegian hospital rolled-out an EHR system to achieve paperless workflows. However, the new EHR system did not initially replace patient (paper-based) records as other information about the patient remained on paper. Thus, in parallel to EHR use, digital records were printed and kept in storage together with any other paper-based documents (e.g. incoming letters). After about three years from the initial EHR roll-out, the organisation decided to address the sustainability issue of this practice by purchasing scanners, including small ones that clinical departments would use to scan paper-based documents to add to the electronic record. This was considered a low-cost 'IT project', 'entailing simple hardware purchase and installation', without the perceived need for changes in workflows or specifically allocated resources. This scanning project began with a pilot of four hospital units, but without an implementation strategy. The researchers were able to observe how this initiative was received by the users in one of these units (the Women's clinic).

From the start, it became apparent that the staff of the clinic encountered several 'sensemaking challenges': how many scanners were needed? Who should scan the documents and when? Should there be changes to the workflow? What should be the processes for handling, distribution, registration, and further processing of documents? And why? Exceptions to initial assumptions surfaced as staff began to explore answers to these questions. The process of making sense took place through tentative plans, meetings, communication with the IT team and other hospitals services, mapping workshops, discussion and negotiation, that progressively involved the entire hospital, over a background of other concurrent change (e.g. the upgrade of the local area network infrastructure).

Overall, 'a seemingly trivial change (the installation of scanners) triggered a larger organisational change process than what had initially been expected.'[16][p24]. Completion of the project took time, but in the end, based on the experiences of the pilot, flexible procedures and organisational standards were developed, workflows were changed, scanning was implemented across all departments and this eventually completed the transition to paperless activities initiated with the initial EHR roll-out.

2.2. How the theory was applied to the case

Aanestad and Jensen took collective mindfulness to be an organisational capability founded on processes of sensemaking. In seeking to understand the hospital implementation of the EHR system, they therefore asked what sensemaking processes occur during post-implementation adaptations and how these processes can be supported for the organisation to achieve this capability. Having had the opportunity to learn about the hospital decision of installing scanners to deal with paper records that remained in use after EHR roll-out, they observed the unfolding of this IT project, participated in meetings and interviewed the staff involved.

They mapped the 'sensemaking challenges' encountered in the hospital by the people who were tasked with making sense of how to adapt their work practices around 'the scanning'. Sensemaking activities were refined in terms of '*making, giving, demanding* and *breaking sense'* [p24] which were then 'translated' into the five processes of collective mindfulness modified to be applicable to 'an action-oriented context'. For example, preoccupation with failure was translated into preoccupation with constraints and preconditions. Deference to expertise, was manifested in the seeking out of an appropriate mix of expertise (Table 3).

In this case study, the collective mindfulness lens reveals the collective cognitive processes and associated activities necessary to bring the initiative to completion hospital wide. The five dimensions are also proposed as an intervention towards future 'successful' implementations. This we discuss in the next section.

Table 3. Dimensions of collective mindfulness in HRO compared to those of a HIT project

Concerns	Collective mindfulness processes in HROs	Collective mindfulness processes in HIT project	Description
Anticipating the unexpected	Preoccupation with failure	Preoccupation with constraints and preconditions	Widespread questioning of preconditions and effects of decisions, seeking to check and validate the assumptions acted upon before decisions were implemented. [16]
	Reluctance to simplify interpretations	Reluctance to premature commitment	Unwillingness to proceed on insufficiently known ground. Decisions and proposals questioned and examined for their upstream and downstream requirements and consequences. Plans considered tentative. [16]
	Sensitivity to operations	Sensitivity to interdependencies and continuous prioritisation	Collectively constructed understanding achieved through collaborative workflow mapping and graphical charting to detect interconnections and dependencies between elements in the work system. [16]
Containing the unexpected	Commitment to resilience	Commitment to avoid disruptions	Maintaining as smooth an operation as possible and minimising disruptive changes as guiding principles in the decision processes. [16]
	Deference to expertise	Seeking out appropriate mix of expertise	In considerations of the effects of planned changes, seeking appropriate mix of expertise - a constellation of actors that would be able to cover the necessary domains and ensure that preconditions and consequences are noticed. [16]

3. Explanations of success or failure of the health IT implementation

Practitioners and researchers of HIT implementations are often concerned with the success or failure of projects. They seek rules or guidelines as 'paths to success', or to prevent or solve HIT implementation failures. What does the theory of collective mindfulness, applied to the case above, contribute to this perspective?

The theory of collective mindfulness originally described activities of HROs as mindful organising enabling the achievement of reliability (prevention and containment of hazards) in high risks businesses. In this light, success and failure pertained to organisations as a whole and reflected the actualisation (or not) of hazards.

Instead, the case explored in this chapter draws attention to the ongoing processes of adoption that make organisational IT evolve and change. From this perspective, the definition of 'implementation success' is rather loose and fuzzy. The hospital eventually reached the goal of 'going paperless' – what we may call 'success' –, but in a more roundabout way than originally envisaged and taking longer than planned. Surprises and unexpected changes are inevitable, given the adaptive complex sociotechnical system where implementation takes place, and the inevitable limitations of all technologies. The case study hospital EHR implementation is not an exception. The theory of collective mindfulness applied to this case gives insights into the ways 'the unexpected' (i.e. the inevitable surprises during implementation) is dealt with, that may turn a problematic implementation process into a 'success'.

In the specific case study described, the organisation demonstrated the capacity to support the discovery of ambiguities, solve emerging issues and progress with an otherwise ambiguous ill-defined project ('the scanning'). The five dimensions of collective mindfulness were present, both to anticipate and contain the unexpected (e.g. issues for others' clinical work generated by changes in workflows). Crucial elements and fundamental premises for collective mindfulness were: widespread questioning of assumptions, of preconditions and of effects of decisions; collective discussion; under-specification of structures; support from project management. These enabled sensemaking activities akin to those practiced in user-centred system design, such as 'co-design sessions' and 'workflow mapping exercises'.

The project team also encountered barriers that made the sensemaking activity more difficult. For example, clinicians were too busy to participate in meetings; expertise in 'their part of the workflow' was missing from the 'appropriate mix' necessary to understand repercussions of proposed changes across the whole process. This was addressed by taking note of necessary questions that one of the team members would ask doctors after the meeting. There was also the inevitable tension of most HIT implementations between standardisation and local customisation. The solution that worked for this hospital was the design of a 'hierarchy of standards': a level 1 hospital-wide procedure, and a level 2 specific to each local department. This structure allowed for a standardised flexibility where local work redesign would take into account interconnections with other departments.

Overall, the case shows how 'mindful managers of change' 'can draw on the five characteristics of mindfulness to ensure more productive organizing to support mindful sensemaking' [16, p26].

4. Discussion and conclusion

The theory of collective mindfulness has been challenged on a number of grounds. HROs as an organisational type remain ill-defined; achieving reliability is not necessarily equivalent to achieving safety; and the five mindful organising principles could be viewed as representing more ideals an organisation may aspire to than a description [17-19]. The five principles are not sufficient on their own for an organisation to operate safely, but need to build on structural preconditions, such as human resources practices that foster trust and respect, and selection and allocation of resources, including IT [20]. The HRO model represents the appropriate organisational response to a certain type of risks and environment, but must not be considered the right response for all environments [21]. Recommendations for practice drawn from the theory may be challenging to put into effect as they involve, for example, changing organisational culture or communication practices.

Despite some criticisms and limitations, the theory of collective mindfulness has been used in healthcare as foundation for informing the development of interventions proposed to improve quality, safety and resilience [18, 22]. However, the theory of collective mindfulness has rarely been applied in health informatics. As shown in the case discussed in this paper, a process of translation may be necessary to make the theory applicable to HIT implementations. Through this act of translation, the case shows how the theory of collective mindfulness can enrich our understanding of organisational processes of adoption and adaptation post-implementation. Indeed, the case also suggests the theory's potential to explain and support the entire HIT project life-cycle – from design, to implementation and evaluation. For example, some of the experiences of the scanning team are akin to those eliciting requirements for system development, or those pre-implementation steps when decisions are taken on what needs to be done, including the necessary changes to workflows. It also shows how theory-based qualitative evaluations of HIT projects based on dimensions of collective mindfulness can explain implementation outcomes.

Healthcare services are high reliability-seeking organisations, struggling to eliminate errors and low-quality patient care. Improving patient safety is the main objective underlying many health IT initiatives – i.e. technology has been endorsed with the key task of helping healthcare organisations achieve safe and reliable patient outcomes. The theory of collective mindfulness explains how HROs achieve reliability of operations. A dimension of these organisations' 'success' is the capability to manage 'the unexpected' despite uncertain and risky conditions. Building on these foundations, we conclude this chapter with a proposal for a twofold definition of success for HIT projects, that others may wish to test:

- HIT is successful when it fosters reliable and safe patient outcomes by sustaining collective mindfulness capabilities.
- HIT implementations are successful not when they avoid 'the unexpected' (perhaps an impossibility) but when they manage the unexpected-related challenges through a mindful *collective mind.*

Teaching questions for reflection

1. What recommendations for HIT practice could be drawn from the application of the theory of collective mindfulness to the implementation of health IT?
2. How can the theory of collective mindfulness assist in evaluating the impact of HIT?
3. What would be the metrics for a mindfulness scale aimed at measuring HIT implementations? And could the scale be used to predict outcomes of HIT project 'success' or 'failure'?
4. How could the theory of collective mindfulness to HIT implementations be translated to consider 'organisations' that are ill-defined, such as in telecare for community care, or in the implementation of patient facing systems, such as a patient portal?

Acknowledgments

Valentina Lichtner received funding from the European Union's Horizon 2020 research and innovation programme under the Marie-Skłodowska Curie Grant Agreement number 740131.

We thank reviewers and editors for their comments to an earlier version of this chapter. Special thanks to Mark Sujan for very insightful and constructive suggestions.

References

[1] E. Langer, *Mindfulness*. 1989: Addison-Wesley.
[2] K.M. Sutcliffe and T.J. Vogus, *Organizing for Mindfulness*, in *The Wiley Blackwell handbook of mindfulness*, A. Ie, C.T. Ngnoumen, and E. Langer, Editors. 2014. p. 407-423.
[3] K.E. Weick, K.M. Sutcliffe, and D. Obstfeld, *Organizing for High Reliability: Processes of Collective Mindfulness.* Research in Organizational Behavior, 1999(21): p. 81-123.
[4] K.E. Weick and K.H. Roberts, *Collective mind in organizations: Heedful interrelating on flight decks.* Administrative science quarterly, 1993: p. 357-381.
[5] K.M. Sutcliffe and K.E. Weick, *Mindful Organising and Resilient Healthcare*, in *Resilient Health Care*, E. Hollnagel, J. Braithwaite, and R. Wears, Editors. 2013, Ashgate: London. p. 145-156.
[6] E. Hutchins, *Cognition in the wild*. 1995, Cambridge, Mass.: MIT Press.
[7] K.E. Weick, *Sensemaking in organizations*. 1995, London: Sage.
[8] K.E. Weick and K.M. Sutcliffe, *Managing the unexpected: Resilience performance in an age of uncertainty*. 2nd ed. 2007, San Francisco, CA: Jossey-Bass.
[9] B.S. Butler and P.H. Gray, *Reliability, Mindfulness, and Information Systems.* MIS Quarterly, 2006. **30**(2): p. 211-224.
[10] T. Aven, *Implications of black swans to the foundations and practice of risk assessment and management.* Reliability Engineering & System Safety, 2015. **134**: p. 83-91.
[11] R.W. Eastburn, *Organizational Mindfulness in Banking: A Discriminating Factor for Firm Performance.* Journal of Managerial Issues, 2018. **30**(1).
[12] J.L. Carlo, K. Lyytinen, and R.J. Boland Jr, *Dialectics of collective minding: Contradictory appropriations of information technology in a high-risk project.* MIS Quarterly, 2012: p. 1081-1108.
[13] T.J. Vogus and K.M. Sutcliffe, *The impact of safety organizing, trusted leadership, and care pathways on reported medication errors in hospital nursing units.* Medical care, 2007. **45**(10): p. 997-1002.
[14] T.J. Vogus and D. Iacobucci, *Creating Highly Reliable Health Care: How Reliability-Enhancing Work Practices Affect Patient Safety in Hospitals.* ILR Review, 2016. **69**(4): p. 911-938.
[15] S. Dernbecher and R. Beck, *The concept of mindfulness in information systems research: a multi-dimensional analysis.* European Journal of Information Systems, 2017. **26**(2): p. 121-142.
[16] M. Aanestad and T.B. Jensen, *Collective mindfulness in post-implementation IS adaptation processes.* Information and Organization, 2016. **26**(1–2): p. 13-27.
[17] A. Hopkins, *The problem of defining high reliability organisations.* National Research Center for Occupational Safety and Health Regulation. January, 2007.
[18] M.A. Sujan, *High Reliability Organisations: Making Care Safer through Reliability and Resilience*, in *Improving Healthcare - A Handbook for Practitioners*, L. Baillie and E. Maxwell, Editors. 2017, Routledge: London. p. 101-114.
[19] N. Leveson, N. Dulac, K. Marais, and J. Carroll, *Moving Beyond Normal Accidents and High Reliability Organizations: A Systems Approach to Safety in Complex Systems.* Organization Studies, 2009. **30**(2-3): p. 227-249.
[20] K.M. Sutcliffe, L. Paine, and P.J. Pronovost, *Re-examining high reliability: actively organising for safety.* BMJ Quality & Safety, 2017. **26**(3): p. 248-251.
[21] R. Amalberti, *Navigating safety: Necessary compromises and trade-offs - Theory and practice*. 2013: Springer.
[22] S. Hines, K. Luna, J. Lofthus, et al., *Becoming a high reliability organization: operational advice for hospital leaders*, in *AHRQ publication*. 2008, Agency for Healthcare Research and Quality: Rockville, MD.

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

doi:10.3233/SHTI190116

Fostering Shared Decision Making with Health Informatics Interventions Based on the Boosting Framework

Marieke DE VRIES[a,b1], Jesse JANSEN[c], Julia VAN WEERT[d] and Rob HOLLAND[b]
[a] *Radboud University, Institute for Computing and Information Sciences (iCIS) Nijmegen, the Netherlands*
[b] *Radboud University, Behavioural Science Institute (BSI), Nijmegen, the Netherlands*
[c] *The University of Sydney, Sydney School of Public Health, Sydney, Australia*
[d] *University of Amsterdam, Amsterdam School of Communication Research/ASCoR, Amsterdam, the Netherlands*

Abstract. The accumulation of medical knowledge, technology and expertise has provided people with more and more options to improve their health and increase longevity. However, healthcare options typically come with benefits as well as harms and often involve important and complex, high-stakes trade-offs. The ideal of Shared Decision Making (SDM), where a healthcare provider and a patient exchange information, bring in their respective professional and existential expertise and consider the options in light of what matters most from the patient's perspective, is a paradigm that is increasingly viewed as a gold standard for high quality care nowadays. eHealth provides ample opportunities to foster personal health choices and SDM through digital information exchange and personal values clarification support. The boosting framework attempts to describe how to foster people's competences to make choices. Its vision is to equip individuals with competences, for instance improved risk literacy, to empower them to make well-informed choices when facing a difficult choice, such as decisions about health issues. Application of the boosting framework to personal health choices and the SDM process unveils new and promising horizons for future research and could inform the design and evaluation of health informatics interventions such as decision support systems.

Keywords. Personal Health Choices, Shared Decision Making (SDM), Decision Psychology, Boosting, Patient Decision Aids (PtDAs)

Learning objectives

After reading this chapter, the reader will be able to:

1. Understand how people's competence to make their own choices can be fostered according to the boosting framework.
2. Understand how the boosting framework can be applied to design and evaluate health decision support interventions, such as patient decision aids.
3. Understand the challenges and opportunities of the boosting framework in the context of health decision support design and evaluation.

[1] Corresponding Author, Marieke de Vries, E-mail: Marieke.deVries@ru.nl.

1. Fostering personal health choices and shared decision making: The boosting framework

Our health and well-being are affected by the choices we make and our ability to act in accordance with those choices. In this chapter, we focus on the competences people need to make and implement personal health choices that align with their values and life goals, and how health informatics interventions can foster those competences. The ultimate aim is to empower people to take a more active role in making choices that can help them to live a healthy and happy life. More specifically, we will focus on shared decision making (SDM): the joint decision-making process through which a patient and his or her healthcare provider exchange information and make a health choice about for example a medical treatment or test [e.g., 1].

SDM is increasingly receiving attention in theorizing about health care practice and is nowadays often viewed as the gold standard, most importantly because of its ethical imperative. However, its implementation in everyday healthcare practice is lagging behind [e.g., 2]. In this chapter, we focus on how health informatics interventions can "boost" the competences of patients and health care providers to make and implement better health choices that align with what matters to the patient. A core assumption in our chapter is that if we want to optimize the effectiveness of health informatics interventions, it is essential to understand the core competences needed to engage in SDM, as well as how to enhance those competences. We focus on a relatively new theoretical framework that is currently gaining traction in psychological science: the boosting framework [3]. Boosting has the potential to inform the design of health informatics interventions by explicating guiding principles for identifying and supporting the competences people need in order to put the widely embraced ideal of SDM into practice.

Boosting aims to empower people in decision making (i.e., to make more beneficial personal choices) by enhancing people's competences and knowledge, based on insights from behavioural science. It is often contrasted with nudging. To set the stage for the boosting framework in the domain of health informatics, we will first describe the concept of nudging before explaining the boosting framework in more detail. Although both approaches intend to change the way in which people behave and make decisions, with the ultimate aim of enhancing their well-being and health, they differ in several ways.

The nudging approach is based on insights from psychology and behavioural economics. With nudging, the choice architecture is changed in order to "nudge" (or gently push) people in the direction of what is considered the "best" option (e.g., the healthiest choice). Nudges change the choice architecture, without changing the reinforcement structure or excluding certain decision options [4]. This way, nudging is assumed to respect an individual's autonomy while making a choice. The nudging approach is based on a series of theoretical assumptions concerning human behaviour and decision making consistent with dual-process models [e.g., 5]. Those dual-process models share the general assumption that even though people are capable of rational and deliberate action, behaviour and decision making are mostly guided by automatic and error-prone psychological processes. Controlling or bypassing these automatic processes is possible, but it requires effort and is therefore considered to be the exception rather than the default process.

Nudging can be appropriate in situations where there is an objectively better option (e.g., quit smoking) and is less appropriate for many personal health care choices, where the "best option" can only be determined based on the physician's medical expertise and the patient's existential expertise. Nudging interventions use automatic processes by triggering automatic tendencies within the choice architecture that steer people towards the "right" direction. A classic example of nudging consists of the facilitation of healthy food choice. People have a strong and automatic preference for high calorie foods and the mere perception of such food items may trigger the automatic urge to buy and eat them. An intervention in canteens based on the nudging approach would place unhealthy products out of sight, even though these products would still be available (i.e., are not banned). Another example is changing the default option. Research has indicated that people tend to go along with the default, because it is easier, and it is pleasant to do what others do. Changing the default option may drastically change people's choices (e.g., the default to donate organs except when actively selecting opt-out). Interventions based on nudging have been widely and successfully used by various companies, organizations and governments. Nudging based Behavioural Insights Teams (BITs; also referred to as "Nudge Units") that assist companies and governments in achieving behaviour change have become increasingly popular the past ten years [e.g., 6].

The boosting framework has been recently introduced by Hertwig and colleagues [3]. It is based on the optimistic view that people are capable of learning new insights and skills. Whereas nudging focuses on changing people's automatic reactions within a specific setting, boosting aims to provide individuals with skills and competences that may help them across situations and time and potentially increases people's sense of autonomy [3]. Both boosting and nudging frameworks are not specific enough to refer to as a theory to explain specific phenomena, because it would entail to make very specific testable predictions in specific situations, whereas boosting and nudging are ways to approach behaviour change interventions. There are some assumptions within the frameworks that could be tested, and we refer to those ideas in the chapter (e.g., about well-being, long term effects), but these predictions are very general and not strong enough to warrant the label theory.

1.1. Theoretical background of boosting

According to the boosting framework, the human mind is malleable. The boosting framework acknowledges the bounds of the human decision maker, such as its vulnerability to cognitive biases and errors in e.g., risk assessment. Yet, it aims to identify existing competences and ways to foster them, for example through improving skills or knowledge, or by providing decision tools. Boosting's view of the human mind is "… that of an adaptive toolbox of ecologically rational heuristics." [3, p. 980].

The theoretical background for boosting is derived from various insights in decision making [e.g., 7]. At its core is the assumption that people's cognitive processes adapt to experiences and that they can rapidly learn to overcome potential errors. Several research traditions provide support for this idea. First, several scholars in decision making argue that human thought is based on experiences and subjective probability. That is, learned patterns may match emerging situations and thereby trigger behaviours that are relevant and useful in that situation [7]. Second, heuristics can be used as tools to make decisions, even under uncertainty. These heuristics help people to make smarter decisions because they are (generally) adaptive within the situation. These approaches largely converge with a Bayesian approach to decision making, in which people use priors (based on previous experiences and/or learning) to predict their environment (e.g., anticipated outcomes and associated experiences in a challenging decision task they are facing) on the basis of available information and resources. In this way people are often able to make smart decisions without elaborate deliberation, using intuitive processes that are based on learned patterns and relevant priors.

Boosts are tools that are based on this optimistic assumption that people have the capacity to learn, and support people in increasing their competence to make good decisions. These boosts can take various shapes, such as simply providing information (e.g., about illness); strategies to make information more easy to understand; tools or skills-training to help clarify the values of decision alternatives (e.g., how to integrate information according to your values); self-management skills (e.g., knowing when and how to act within treatment or revalidation; knowing how to make implementation intentions); social skills (e.g., how to approach another person during an important conversation) and others. These boosts can be domain specific (e.g., health information about a certain illness) or more general (e.g., strategies to improve statistical literacy). As we will explain in more detail in section 2, boosts can improve health related decisions. For example, these boosts increase competence by helping patients to understand risk information or to make decisions in line with one's core values.

1.2. Autonomy and well-being

Boosts make people more competent. Therefore, by definition, boosts are beneficial for the individual. In addition, increased competences have positive effects on psychological processes including motivation, autonomy and well-being. For example, improved knowledge enhances self-efficacy, i.e., the core belief in the ability to achieve self-relevant goals. Subsequently, people with high efficacy may display stronger motivation; e.g., persist longer in the face of obstacles [8]. More generally, this way, boosts may increase feelings of autonomy and well-being.

In contrast, autonomy is not fully respected in nudging. People are often not aware that they are being nudged. Instead, they depend on the good intentions of the "nudger", e.g., healthcare providers, the government, institutions and companies, acting as a so-called "benevolent dictator". The goals of nudgers are not necessarily consistent with the goals of the decision maker. This is crystal clear when nudging is used for commercial purposes, but even when the government or healthcare providers operate as nudgers, goals may not converge with the values of an individual, and hence, may have a negative impact on the well-being of those being nudged. Therefore, nudging may be considered the best approach only when 1) there is great consensus among individuals within society concerning the necessity of behaviour change and 2) individuals are not motivated or able to learn skills to change this behaviour [9].

When the values or preferences of individuals are highly heterogeneous and good choices fully depend on people's own values, boosting clearly outperforms nudging. In cases in which the values of the person are not known, nudging is like playing roulette, whereas boosting could help people to integrate their own values in a choice. Thus, as Hertwig [9] indicates, when individual values are at stake, boosting is strongly preferred to nudging. This is often the case in health-related decision making, where there is likely to be a balance between the benefits and harms of different options (decisional equipoise) and/or decisions are preference sensitive because of the variation of how people value attributes of different options. People may have strong feelings about decisions where the best option seems obvious, e.g., removing the contralateral breast in women with average breast cancer has no mortality benefit but an individual woman may still prefer it because she wants to feel balanced. Some scholars argue that all health decisions are potentially preference sensitive, implying that physicians should always consider the patient's perspective, even in cases of decisions about issues that may appear value-neutral at first glance [10].

Nudges are (mostly) restricted within the manipulated situation. That is, within the framework, there are no assumed spill-over effects of changes in one choice architecture to choices in other contexts. Even though nudges may potentially have a long-lasting effect within the manipulated context, such stable effects are rare, and unlikely as the behaviour will go back to the default mode when the nudge is removed from the situation.

1.3. Shared decision making (SDM)

Modern-day health care practice offers people an increasing amount of healthcare options, from preconception to end-of-life care. Advances in medical knowledge and health innovation have also resulted in increasingly complex decisions regarding personal health. Most (if not all) of us face some through personal health choices throughout our life courses. Not only do (preference-sensitive) medical decisions often involve high-stakes and trade-offs between potential benefits and harms of different options (such as between quantity and quality of life, or between treatment efficacy and treatment burden), as patients we also often face emotionally charged, unanticipated and novel situations. To optimize personal health choices, new information needs to be integrated with personal values, life goals and circumstances.

People vary widely in how they value the matters at stake in trade-offs. For each individual patient, the suitability of each of the medical treatment options depends on the individual patient's unique values, preferences and circumstances. This makes it essential to involve patients in the decisions that concern their life and well-being, as patients also indicate themselves – about 80% of people want to be actively involved in the medical decision-making process involving invasive medical procedures [11]. Another example that illustrates the need for more active patient engagement in treatment decision making comes from a recent study in the Netherlands, which revealed that one in three prostate cancer patients was dissatisfied with the amount of information they received about their treatment options [12]. Shared decision making is not only aimed at physical health, but also applies to mental health [13].

SDM is often characterized as a meeting between two experts: A medical and an existential expert. The medical expert, that is, the healthcare provider, can bring in professional expertise, such as information about the medical condition from which a patient is suffering, the medical treatment options for which a patient is eligible and the evidence about the pros and cons associated with those options, according to available medical evidence and the healthcare provider's own expertise. The existential expert, that is, the patient (and in some cases, such as in aged care, also their loved ones), can bring in information about his or her unique circumstances, personal values and (life) goals, which are essential for interpreting the medical evidence in light of what matters most from the patient's perspective. Even if a patient does not want to make a final decision, SDM can help a health care provider to make a decision that is sensitive to the patient's values and context. The ideal of SDM has been called "the pinnacle of patient-centered care" [14]. However, despite SDM being embraced as the gold standard, its implementation is lagging behind and health care practice still widely deviates from this norm. For example, in a study published in 2012, less than 50% of patients reported that their healthcare providers had considered their personal goals or concerns [15].

1.4. Using health informatics to foster personal health choices and the SDM process: The case of patient decision aids

One clear example where boosting applies to health informatics is the design of patient decision aids (PtDAs) aimed at supporting personal health choices and SDM. PtDAs are interventions that “support patients by making their decisions explicit, providing evidence-based information about options and associated benefits/harms, and helping clarify congruence between decisions and personal values” [16, page 1]. PtDAs can be used in preparation for the visit with a clinician, during the visit or individually by the patient, for example in the context of breast screening decisions. When used without input from the clinician, PtDAs aim to support informed choice rather than SDM per se. Even though, to the best of our knowledge, there is no existing case of a health informatics intervention aimed at fostering SDM that explicitly used the boosting framework to inform its design, many existing ways to foster SDM are consistent with the boosting framework. Throughout the remainder of this chapter, we will explicate how the boosting framework applies to existing cases and could be used to further inform future design of health informatics interventions aimed at fostering SDM.

In the eHealth era, information technology provides ample opportunities to unlock and share valuable information resources, such as information exchange and supporting patients and their healthcare providers in making well-informed medical decisions that align with what matters most for the person whose values are at stake: the patient. In other words, health informatics has the potential to boost decision making capacity. However, in order to be effective, health informatics interventions need to be well attuned to the way the human mind is wired and to the way the care process takes place. The boosting framework can support the design of health informatics interventions such as tailored text messages, online health information tools and PtDAs. Experimental research showed for instance that messages that were personalized (tailored) to the individual (“boosts”) led to a higher decrease in snacking consumption than non-tailored messages [17]. Research has also revealed that online health information tools are facilitating immediate, intermediate and long-term (including clinical) patient outcomes, even in older patients. In particular those tools that not only provide information, but also have self-management and/or information exchange functions, exactly the functions that can serve as boosts, seemed to be effective [18]. Although the majority of existing PtDAs are paper-based, not yet digital [16], the development of online PtDAs is rapidly increasing. Using online PtDAs has several benefits, including the possibility to provide personalized information, tailored to individual patient information needs and to be more interactive. Finally, the use of “big data” in patient and provider decision support allows to access and use vast amounts of data that have been collected for other purposes (such as cancer registry data) but may be valuable in the SDM context as well [19]. Recent research uses for instance prediction models based on “big data” to estimate personalized risks and outcomes, such as drug interactions, or treatment (side) effects.

For the design and evaluation of PtDAs, it is important to build on insights from relevant theories such as the boosting framework, to ensure that decision support interventions help their users to harness their decision pitfalls and to foster their decision competences [20]. This is especially important if a PtDA is intended to be used by a patient without the input of a clinician (e.g., to support cancer screening decisions) In section 1.5, we elaborate on the design of health informatics interventions which aim to inform patients about eligible healthcare options. In section 1.6, we elaborate on the design of values clarification methods in health informatics interventions.

1.5. Medical information provision through patient decision aids

In order to make well-informed medical decisions, the patient and the healthcare provider need access to medical information about the options to which a patient is eligible, including risk information. The vision behind the boosting framework aligns with this need: to equip individuals with competences, such as risk literacy competences, and hence, to empower them to make well-informed choices when facing a difficult personal choice.

A classic example of how the boosting framework may help to make better decisions and can inform the design of more effective ways to communicate risk information involves boosts that help to better understand medical risk information. Generally, patients as well as healthcare providers have difficulty understanding conditional probabilities [e.g., 21]. For instance, consider the conditional probabilities for breast cancer. Let's assume the base-rate (prevalence) to get breast cancer is one out of 100 women. The accuracy of a mammogram, an X-ray test to indicate whether a person has breast cancer, is about 80-90%. More specifically, the probability of the mammogram resulting in a positive test result when breast cancer is present (sensitivity) is 80%. The probability of the test result of the mammogram being negative when the disease is absent (specificity) is 90%. Now, a woman is tested positively on the X-ray test, what is the chance this woman has breast cancer? In other words, what is the positive predictive value of the mammogram, what is the probability that a patient has the disease when the test result is positive? Both healthcare providers and patients typically overestimate this chance and judge it to be around 75%, whereas the actual chance is much lower: It is only 7-8%. This is because people tend to neglect the base-rate. Gigerenzer and colleagues indicated that we could make these risks much more understandable for healthcare providers and patients by using natural frequencies rather than conditional probabilities. Risk information about breast cancer would then be explained in the following way: Out of 1000 women, 10 women will have breast cancer and 990 will not. Out of those 10 women who *do* have breast cancer, 9 will receive a positive result on the X-ray test and 1 will not (false negative). Out of the 990 women who *do not* have breast cancer, 99 will receive a positive result (false positive) and 891 women will receive a negative result. This way, it is more transparent to see the role of base-rates: many women *without* cancer are in fact *tested* positively. This approach to presenting risk information can be considered a boost, because by presenting risk information in terms of frequency information, the understanding of the information increases and therefore potentially the quality of decisions based on this information increases as well. In a similar way, illustrations, animations and videos can serve as boosts. Illustrations, in particular those supporting a text, are widely used to facilitate learning of information by improving comprehension and recall [22]. Adding videos to online texts, particularly personalized videos using a conversational narration style, also improves memory for medical information [22], and animations can even bridge the information processing gap between audiences with low and high health literacy [23].

1.6. Values clarification though patient decision aids

For the alignment of medical decisions to an individual patient's (often implicit) values and personal circumstances, patients need to clarify their personal values and preferences. This can be challenging. Moreover, potential outcomes and risks associated with the medical options to which a patient is eligible, can be hard to imagine or hard to verbalize, and the available options often involve important, high-stakes and highly personal trade-offs, such as those between quality and quantity of life, which cannot be solved in a straightforward manner. How much future quantity of life would you be willing to "trade" for a better quality of life right now, for example? Taking people's values into account is even more important for specific groups at risk, such as people with multimorbidities or older people, for whom no or limited clinical evidence is available. Guidelines for treatments are usually based on studies in which those groups were excluded [23], resulting in a lack of detailed information about the optimal treatment.

Some PtDAs do not only provide information, but also include additional content aimed at supporting patients to clarify their personal values and preferences: Values Clarification Methods (VCMs). Every tool that provides patients better insights about their values can be considered a boost, as these tools make participants more competent in processing and weighing their values. From a boosting perspective [3] it is important to systematically analyze which competences patients (and/or healthcare providers) are naturally possessing or lacking in this regard and to create tools (boosts) to augment or overcome these.

2. Explanation of success or failure in health informatics interventions for SDM

2.1. Success factors and failures of patient decision aid design and evaluation

The design and evaluation of most PtDAs is heavily informed by the International Patient Decision Aid Standards [IPDAS; 24]. The IPDAS collaboration is a group of international researchers, practitioners and stakeholders who have outlined a systematic process for PtDA development and evaluation, as well as specific recommendations, e.g., information presentation and values clarification methods (VCMs). It comes with a set of quality criteria and reporting standards to help ensure that PtDAs are of high quality, accurate and unbiased. This is essential because PtDAs can have an important influence on decisions made [16]. The IPDAS quality criteria are related to the following dimensions: 1) information provision, 2) presentation of outcome probabilities (risk communication), 3) clarifying and expressing values (VCMs), 4) decision guidance, 5) using a systematic development process, 6) using evidence, 7) disclosure and transparency (COI), 8) use of plain language, 9) and evaluation of PtDA effectiveness. The development process to a large extent builds on the Ottawa Decision Support Framework (ODSF) which is guided by expectancy value, decisional conflict, and social support theories, but is mostly consensus based [24]. It describes PtDA development as an iterative process which includes extensive involvement of and testing with patients and healthcare providers. The IPDAS recommendations and quality criteria draw on systematic reviews of available evidence, including those for information provision and risk communication and values clarification. These recommendations are currently being updated.

The state-of-the-art knowledge about the effectiveness of PtDAs to support people in making medical treatment or screening decisions is promising: an accumulating amount of research, including over 100 PtDAs studies across a variety of medical treatment and screening decisions, has showed that patient decision aids have been effective in improving people's knowledge, feelings of being well-informed and clarity about their personal values. The evidence is less clear with regard to other outcomes, but it appears likely that people who have been exposed to PtDAs also have more accurate knowledge of the benefits and harms associated with medical options and have been more actively involved in the decision making process. Also, there is some evidence suggesting that PtDAs may help people to make choices that are congruent with their personal values and preferences, but more research is needed in this area. PtDAs do not have adverse effects on health outcomes or patient satisfaction. More research is needed to determine if PtDAs help people receive and adhere to their chosen option [16] and if this results in, for example, better well-being and quality of life.

Even though the available evidence shows that the use of PtDAs has the potential to help people make better health decisions (in terms of better matching their values and preferences), more fine-grained understanding of the underlying processes is limited. Most studies so far have compared PtDAs to usual care rather than using a study design suitable to identify "active ingredients". This limits the extent to which evidence-based guidance for designing effective PtDAs can be formulated. A theory-based approach can help to further the field. In the past, VCMs for example have often been designed without clearly being rooted in theory [25]. This may have hampered VCM effectiveness and complicated transparent and coherent, systematic design and evaluation. Where theory has been used, debate followed about the appropriateness of theory for the design of VCMs [e.g., 20,26]. Similarly, there is a gap in the evidence base for risk communication in the context of PtDAs. For example, it is still unclear how we can tailor risk communication in the context of eHealth and interactive tools to individual needs and abilities. What are optimal risk communication formats for vulnerable groups, including those with lower health literacy, numeracy and/or graph literacy? Rooting the design of PtDAs (including VCMs and risk communication) in theory and providing a clear rationale on how the theory has informed PtDAs design, enables targeted tests of the underlying mechanisms and may ultimately help to uncover the "active ingredients" of PtDAs. Addressing this gap will require systematic testing of different information formats within the same PtDA.

2.2. Challenges and opportunities of the boosting framework in the context of health decision support design and evaluation.

Application of the boosting framework to personal health choices and the SDM process unveils new and promising horizons for future research and could inform the design and evaluation of health informatics interventions aimed at facilitating SDM, including PtDAs. In this section, we elaborate on the challenges and opportunities of the boosting framework in the context of health decision support design and evaluation.

The main opportunities of using the boosting framework in this context are the guidance it can provide to design and test health informatics interventions that fit the way the human mind is wired, so that the interventions are likely to be more user-friendly, useful and effective–and therefore also more likely to be implemented and used sustainably. Implementation of PtDAs, and more broadly speaking of SDM, in everyday healthcare has so far been a major challenge [e.g., 2]. In a recent study [27] investigating usage of a patient decision aid, observing a sample of more than 1000 patients diagnosed with prostate cancer, only about one in three eligible patients received a link to an online patient decision aid. Those who did receive a link to the decision aid, typically also accessed the decision aid online, utilized most of its content and functions, and discussed the decision aid summary in a follow-up consultation with their health care provider. Even though the overall implementation rate was low in this study, a wide variation in implementation rate (16-84%) was observed between hospitals. Even though the boosting framework may not be sufficient in overcoming the implementation challenge, it may provide some useful opportunities.

The main challenges of using the boosting framework in the context of fostering personal health choices and SDM with health informatics interventions, may very well exist in implicit or explicit resistance to adopting new roles in the patient/clinician encounter, on the part of patients as well as healthcare providers. For example, many patients do not dare to voice their preferences, needs or concerns, out of fear of being labelled a "difficult" patient [e.g., 28]. Healthcare providers may mistakenly assume that their patients do not want to or are not able to take a more active role, but research shows the contrary [18]. Healthcare providers may also have false beliefs about the amount of extra time needed for SDM: Even though the effect of using PtDAs on consultation length is typically about 2 to 3 minutes [16], the belief that SDM is too time consuming is one of the main clinician-reported barriers to implementing SDM in everyday clinical practice[29].

The boosting framework can help shed light on promising avenues for future research. Ultimately, the aim of SDM is to help patients and caregivers make well-informed decisions in collaboration with the health care provider and aligning with what matters most from the patient's perspective. Whereas it is by now well established that PtDAs can help ("boost") people's competences to understand the medical information relevant to the medical choice they are facing and to clarify their personal values, far less is known about how to help people implement their preferences in everyday healthcare. We believe the vision behind the boosting framework unveils new and promising horizons for future research. Boosting is focused on competences people need to make better decisions when they face a challenging decision. Competences that have so far been "boosted" in the field of SDM-focused health informatics, are mainly related to understanding (risk) information and values clarification. In the clinical encounter, where patients and healthcare providers implement the final health decisions, other crucial competences are at play as well, which may very well lie on the social, interpersonal dimension. For example, certain decisions may mean deviating from clinical guidelines and this requires courage, trust and tolerance of uncertainty [e.g., 30]. If the ultimate aim is to empower people so that they can make choices which result in tailored care that truly aligns with what matters most to them, we may therefore need to shift gears, broaden the scope and focus on boosting those other competences that may very well be crucial in driving the ultimate SDM behaviour in the clinical encounter.

3. Discussion

In this section, we reflect on the value of the boosting framework in the context of personal health choices and SDM and on the maturity of the boosting framework in this context.

Generally speaking, the use of theory and frameworks in the design and evaluation of health informatics interventions aimed at supporting health choices and SDM comes with certain advantages compared to building health informatics interventions based on common sense. First, theories and frameworks are more consistent with the state-of-the art scientific knowledge and facts, in this case mainly from the field of decision-making psychology, than common sense. This makes theory-based design more likely to result in effective interventions. Moreover, a theoretical framework, such as the boosting framework, can be developed into a more explicit process theory tailored to the field of personal health choices and SDM, from which testable assumptions can be derived. By empirically testing these assumptions, the theory matures, with adaptions based on empirical findings where necessary, which can in turn yield new testable assumptions.

Currently, the boosting framework does not yet provide a full-blown process model with detailed "how-to" information describing how research evidence can be translated into practical health informatics solutions. Rather, we believe the boosting framework helps to explicate some guiding principles for future research, from which testable assumptions can be derived. To find the most promising avenues for future research, we should start with the end in mind and stay focused on the ultimate aim: Helping patients and caregivers to make well-informed medical decisions that align with what matters most from the patient's personal perspective [31]. Table 1 provides some examples to illustrate the potential of the boosting framework in the SDM context.

Table 1. Health Informatics in SDM: Guiding principles and testable assumptions derived from the boosting framework.

Guiding Principle	Testable Assumptions	Empirical Evidence
Competences can help people make better decisions; boosts can exist of making information easier to understand; or training (more difficult) skills	(1) Presenting risk information in natural frequencies improves understanding (2) Adding illustrations, animations and videos to text-based information improves understanding (3) Training social skills will smoothen interaction between patient and health care provider	Partly available
Decisions based on acquired competences increase people's autonomy	(1) Decisions based on boosts increase autonomy (2) More adherence to patient preferences, (3) More well-being (specifically in the long run)	More research needed
Competences can be used across different patient populations, situations and time	(1) People can acquire decision-making competences relevant to SDM through formal education (2) Healthcare providers can acquire generic SDM competences in professional education	More research needed

By systematically testing these assumptions a process model for boosting SDM can be developed, which can inform the design and evaluation of future health informatics interventions, aimed at boosting all the crucial competences people need in order to be able to make personal health choices that truly line up with their key values and serve their health and well-being in the long run.

The boosting framework is ideally suited to inform the development of health informatics interventions where patients have a choice. By building on existing competences and supporting learning, it has the potential to support autonomy and empower patients to take a more active role in making a decision that is informed and in line with their personal preferences and values. Boosting also highlights the importance of tailoring interventions and the intervention context based on what we know about psychological processes. This is an area in need of systematic research, comparing the effect of different methods of information provision and values elicitation on a broad range of outcomes. The ultimate goal of any decision support intervention is to support patients and their healthcare providers in making evidence-based, informed decisions that are in line with a patient's personal values and preferences. The boosting framework might help achieve this in a way that maximizes patient autonomy while at the same time reducing decisional burden.

Teaching questions for reflection

1. What do designers of health informatics interventions aimed at fostering SDM need to know about the theoretical approach of boosting?
2. What do designers of health informatics interventions aimed at fostering SDM need to know about the similarities and differences between the boosting and the nudging framework for supporting human decision makers?
3. How likely is it that basing the design of a health informatics intervention aimed at fostering SDM (e.g., a PtDA) will lead to improved SDM? Why?
4. What is needed to help the field of health informatics move forward and understand how the vision behind the boosting framework can be applied to have the strongest impact on fostering SDM?

References

[1] C. Charles et al., Shared decision-making in the medical encounter: What does it mean? (or it takes two to tango), *Social Science & Medicine* **44** (1996), 681-692.
[2] A.M. Stiggelbout et al., Shared decision making: Really putting patients at the centre of healthcare. *BMJ* **344** (2013), e256.
[3] R. Hertwig & T. Grune-Yanoff, Nudging and Boosting: Steering or empowering good decisions. *Perspectives on Psychological Science* **12** (2017), 973-986.
[4] R. Thaler & C.R. Sunstein. *Nudge: Improving Decisions About Health, Wealth and Happiness*, Simon & Schuster, New York, 2008.
[5] J.W. Sherman et al. (Eds.), *Dual-Process Theories of the Human Mind*. Guilford Press, New York, 2014.
[6] R. Hertwig, When to consider boosting: Some rules for policy-makers. *Behavioural Public Policy* **1** (2017), 143-161.
[7] G. Gigerenzer et al., *Simple Heuristics That Make Us Smart,* Oxford University Press, New York, 1999.
[8] G. Klein, Naturalistic decision making, *Human Factors* **50** (2008), 456-460.
[9] M. Oaksford & N. Chater, *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*. Oxford University Press, New York, 2009.
[10] A.A. Kon, The shared-decision making continuum. *JAMA* **304** (2010), 903-904.

[11] D.J. Mazur & D.H. Hickam, Patients' preferences for risk disclosure and role in decision making for invasive medical procedures. *Journal of General Internal Medicine* **12** (1997), 114-117.
[12] R.E.D. Lamers et al. Patients are dissatisfied with information provision: Perceived information provision and quality of life in prostate cancer patients. *Psycho-Oncology* **25** (2016), 633-640.
[13] M. Slade, Implementing shared decision making in routine mental health care, *World Psychiatry* **16** (2017), 146-153.
[14] M.J. Barry & S. Edeman-Levitan, Shared decision making - the pinnacle of patient-centered care, *New England Journal of Medicine* **366** (2012), 780-781.
[15] Alston et al., *Communicating with Patients on Health Care Evidence*, Institute of Medicine, Washington, 2012.
[16] D. Stacey et al., Decision aids for people facing health treatment or screening decisions. *Cochrane Database of Systematic Reviews* **4** (2017).
[17] M. Kaptein, Adaptive persuasive systems: A study of tailored persuasive text messages to reduce snacking, *ACM Transactions on Interactive Intelligent Systems (TiiS)* **2**, 10.
[18] S. Bolle, Online health information tool effectiveness for older patients: A systematic review of the literature. *Journal of Health Communication* **20** (2015), 1067-1083.
[19] R. Vromans et al., Communicating personalized cancer statistics: Challenges and opportunities, Abstract from *Etmaal van de Communicatiewetenschap* (2019).
[20] M. de Vries et al., Combining deliberation and intuition in patient decision support. *Patient Education and Counseling **91*** (2013), *154-160.*
[21] S. Plous, *The Psychology of Judgment and Decision Making*, McGraw-Hill, New York, 1993.
[22] F.E. Stubenrouch et al., A web-based application to communicate benefits and risks of surgical treatments. *Surgical Technology International* **30** (2017), 31-37.
[23] I.A. Scott & G.H. Guyatt, Cautionary tales in the interpretation of clinical studies involving older persons, *Archives of Internal Medicine* **170** (2010), 587-595.
[24] A. Coulter et al., (2013). A systematic development process for patient decision aids. *BMC Medical Informatics and Decision Making* **13** (2013), S2.
[25] A.H. Pieterse et al., Theory-informed design of values clarification methods: A cognitive psychological perspective on patient treatment decision making, *Social Science & Medicine* **77** (2013), 156-163.
[26] A.H. Pieterse & M. De Vries, On the suitability of fast and frugal heuristics for designing values clarification methods in patient decision aids: A critical analysis. *Health Expectations* **16** (2013), e73-e79.
[27] M. Cuypers, Uptake and usage of an online prostate cancer treatment decision aid in Dutch clinical practice: A quantitative analysis from the Prostate Cancer Patient Centered Care trial. *Health Informatics Journal* (2018), 1460458218779110.
[28] D.L. Frosch et al., Authoritarian physicians and patients' fear of being labeled "difficult" among key obstacles to shared decision making, *Health Affairs* **31** (2012), 1030-1038.
[29] F. Légaré et al., Barriers and facilitators to implementing shared decision-making in clinical practice: Update of a systematic review of health professionals' perceptions, *Patient Education and Counseling*, **73** (2008), 526-535.
[30] J. Jansen et al. Too much medicine in older people? Deprescribing through shared decision making, *BMJ,* **353** (2016), i2893.
[31] M.J. Barry et al., Shared decision making: Staying focused on the ultimate goal, *Catalyst NEJM.*(2018).

P. Scott et al. (Eds.)

doi:10.3233/SHTI190117

Junior Doctor Communication Systems and the Deterioration Communication Management Theory

Janet LIANG [a,1], Jim WARREN[a], Martin ORR[b] and Karen DAY[a]
[a] *The University of Auckland Private Bag 92019, Auckland 1142, New Zealand*
[b] *AUT University Private Bag 92006, Auckland 1142, New Zealand*

Abstract. Inadequate communication is a factor in suboptimal junior doctor management of deteriorating ward patients. Junior doctors' information and communication technology (ICT) systems are not the sole cause or cure for this. However, junior doctors are already dissatisfied with existing technologies for general hospital communication. The Deterioration Communication Management Theory (DCMT) provides a means to approach these issues by uniting two themes: 1) factors affecting the properties of ICT used to communicate to junior doctors; and 2) factors affecting junior doctor interpretation of communication about deteriorating hospital patients. ICT factors include how the combination of physical devices and mode of usage affect user perception of system reliability and efficiency. Junior doctors interpret clinician communication about patient deterioration in terms of risk, which is affected by their contextual responsibility and experience. Perceived risk and contextual experience in turn affects their communication efficiency. Combining these themes gives more options to explain junior doctor communication in this clinical context and to design ICT systems to improve it.

Keywords. Clinical deterioration; Smartphone; Text messaging, Patient safety, Hospital communication systems

Learning objectives

After reading the chapter, the reader should be able to:

1. Understand the drivers for improvement of hospital junior doctor communication systems, both in general use and in the context of the deteriorating ward patient.
2. Critique the present status of hospital organizational level interventions to improve clinician communication about deteriorating ward patients.
3. Compare different theories used for hospital communication systems and especially those suitable for group communication.
4. Interpret how communication systems can be applied within a hospital setting, and contribute to the design of new systems.

[1] Corresponding Author: Janet Liang, E-mail: jlia002@aucklanduni.ac.nz

1. The drivers for improvement of junior doctor communication systems in the context of the deteriorating ward patient

Deficiencies exist in clinician detection and management of the deteriorating hospital ward patient despite organisational and educational interventions [1], which may be addressed by improving inter-clinician communication. Deterioration in this sense refers to life-threatening or potentially life-threatening changes. Nurses are usually the first hospital clinicians to detect deterioration via observation and measuring physiological changes. They then notify doctors, usually juniors on the ward team primarily responsible for that individual patient's hospital care. These doctors are important in initiating investigations, therapies and escalating patients to clinicians with greater experience or resources to treat deteriorating patients. Nursing or medical mistakes in this process can and do result in adverse patient outcomes [2].

Rapid response systems (RRS) are hospital level interventions which have been widely adopted internationally to improve the process of detection, clinician communication and clinical management [1]. They consist of afferent and efferent limbs, and an administrative component. Afferent limbs function mostly by defining thresholds for single or grouped physiological observations to trigger closer patient observation and clinician communication. Efferent limbs are the triggered clinician responses, ranging from attendance of ward doctors and nurses to a formalised group of more experienced ward and specialty area clinicians, e.g. medical emergency teams. This may culminate in patient transfer to higher acuity hospital units for the care of highly ill patients, e.g. intensive care units. Despite widespread RRS adoption, there is mixed evidence regarding RRS efficacy in decreasing adverse patient outcomes [1]. RRS afferent limb sensitivity is poor unless high false negative rates are accepted. This increases clinician work by requiring closer observation which does not change management [3]. There is also controversy about how well physiological observations in themselves identify deteriorating patients [4] assuming that they are correctly documented [5].

RRS efferent limbs via medical emergency teams may also cause unintended and counterproductive effects on clinician communication e.g. antagonism between ward and medical emergency teams [6]. This situation has now been complicated further by electronic alerting systems, often based on RRSs. These attempt to overcome some of the difficulties encountered in collecting and disseminating patient data for the afferent limb [7]. Research on electronic alerting systems is heterogenous and has shown mixed results. In addition, electronic alerting system studies vary in what criteria were used to trigger alerts, which clinician receives alerts and how [7]. Consequently, it cannot be said that RRS is an entirely benign intervention, or that electronic alerting systems clearly improve the detection and management of the deteriorating ward patient.

Focusing on junior doctors in particular, there are multiple reasons why they may not detect and manage deteriorating ward patients appropriately [8]. Despite undergraduate medical education interventions [9], junior doctors still often feel unprepared working in hospitals, especially with acutely unwell patients [10]. They need to adapt their theoretical knowledge to clinical environments and medical hierarchies, as well as being able to do procedural skills [11]. Inter-clinician communication is also an area that junior doctors find difficult. However, there has been little investigation into general junior doctor communication prior, parallel or alternative to RRS activation [12]. There is widespread ward non-compliance with RRS afferent limbs, including delayed escalation of deteriorating ward patients [13]. It is not just the most junior doctors who may be hesitant to warn senior doctors of patient deterioration [11]; more experienced junior doctors may not do so either [14]. Attempts have been made to standardize and improve a specific communication genre that is likely to occur during a deteriorating patient's course, i.e. handover [15]. Handover (also called handoff) is defined as the process of information, authority and responsibility transfer for a patient from one clinician to another [16]. It is difficult to know the effects of handover interventions on patient outcome [17]. More research is required about why junior doctors communicate as they do.

How and when junior doctors receive and send communication about the deteriorating patient is influenced by both their clinical roles and their information and communication technology (ICT) [18]. For instance, deteriorating patient communications are a relatively small part of a junior ward doctor's large total communication load and need to be differentiated from routine messages [19]. Junior doctor specialty will also affect communication load [20]. Message prioritisation and allocation can be achieved technologically, although clinicians can and will circumvent this if it does not suit their personal purposes [19]. Individuals may also configure communication technologies to suit them e.g. their mobile phones. Where communication technology problems exceed the individual and span an organisation, this may be addressed by changing how the technology is used rather than by changing the technology in itself [21]. Although junior doctors prefer communication technology which is easy to use and increases their personal efficiency, it is unclear which physical device and software combination will achieve this [22,23]. Junior doctors do not have a consistent preference for any one hospital communication system, although generally they are dissatisfied with locators and prefer mobile phones [23]. Closer examination of junior doctor mobile phone usage reveals more complex preferences. Examples include preferences for different communication software in different contexts, and different preferences for communication devices between message senders and receivers [23]. A single device and software combination that addresses all requirements may be impossible, and if technological combinations are used this will likely evolve with time and context [24]. The usage of ICT is thus nested within hospital systems [25]. Consequently, the term information and communication technology (ICT) system is used here to reflect a broad sociotechnical interpretation of technology and context. Software mode is used to refer to a group of software programs sharing similar communicative properties separate to the physical devices hosting them, e.g. task management systems [22]. Whilst physical device types may also share properties e.g. landline and mobile phones, they often have more dissimilarities in other features like mobility and configurability. Either way, both the physical device [26] or the software sending mode [23] can affect how clinicians interpret messages.

Pre-existing communication and communication technology theories do not wholly address the above issues. Communication challenges do exist elsewhere (e.g. the airline industry) although healthcare has specific communication characteristics and needs [27]. The healthcare domain has high cognitive demands from information complexity, unpredictable interaction within different specialties and frequent transfer of responsibility [28]. Morrow and Lopez [28] highlighted information processing, persuasion and risk communication, communication as interaction and common ground theories as being relevant to healthcare communication. These theories have not been used in the specific context of junior doctor communication about deteriorating ward patients. Theories concerning junior doctor usage of ICT systems are also limited in that they may only consider a single device or mode, rather than the multiple choices available at one time. Present ICT adoption theories also may not consider contextual influences on ICT system device or mode selection e.g. message receipt at different time of day or concurrent activities.

The focus of most hospital communication technology intervention studies has been on quantifiable effects on communication metrics and clinician satisfaction, with the minority of studies looking at sociological impacts [22]. The latter study types can demonstrate how complex clinician ICT system requirements may not be well represented by an overall satisfaction metric. An example is Johnston et al [29] where clinicians agreed that electronic communication required adequate security systems. This can conflict with simultaneous demands for optimal efficiency [18], and rapid and accessible inter-doctor communication [30]. Consequently, there is precedent to research junior doctor communication. Deterioration Communication Management Theory (DCMT) [18] diverges from the RRS paradigm by explicitly focusing on junior doctor communication. By doing this, the intent is to inform the design and implementation of organisational and technological interventions to improve the care of deteriorating hospital patients.

2. Introduction to the Deterioration Communication Management Theory

The purpose of DCMT is to improve the design and implementation of ICT systems for communication to and from junior hospital doctors, especially in situations where there is a deteriorating or potentially deteriorating ward patient. ICT system factors are not the sole cause of communication deficiencies in the latter context, although they add to its complexity. Nurses mostly send non-urgent messages so urgent messages about deteriorating patients need to be easy to prioritise. Clinicians may use more than one system during a communication episode. However, they may disagree on which system they prefer for which task, thereby risking missed messages. At the same time, ICT systems can be designed to improve information transmission, e.g. by enforcing better message content and structure. DCMT may also assist understanding where similar ICT system implementations are received differently at different hospitals [23].

DCMT is similar to Stewart's [31] grounded theory of junior hospital doctors' decision making regarding calling senior doctors for help on any matter. Where DCMT diverges is in that it concerns communication decision making by all doctors below the senior level. DCMT is also different in that it specifically addresses communication decision making about the deteriorating ward patient involving ICT systems. It does not encompass the enormity of hospital and medical team [32] communication at large. Rather, as a classical grounded theory is intended to do [33], DCMT is a substantive theory that is meant to explain and account for a pattern of behaviour i.e. how a junior doctor decides whether to communicate about a potential deteriorating patient.

DCMT combines research about clinician usage of hospital ICT systems and links this to recipient junior doctor response to ensuing communication about deteriorating ward patients. It was derived from literature reviews and New Zealand hospital-based studies. The literature reviews [18] were of clinical ICT systems used by junior hospital doctors, and of interventions to improve the detection and management of the deteriorating ward patient. Inter-clinician communication about deteriorating ward patients and ICT systems used in this clinical context were investigated at the primary study hospital [18]. This included pre-existing ICT systems in the primary study hospital such as locators, landline and mobile phones, and two new task manager systems implemented during the research period. The developers of two other task manager systems in other New Zealand hospitals were also interviewed [18]. This was because of potential implementation at the primary study hospital. The remainder of this section is an explanation of DCMT [18] as displayed in Figure 1 below.

Factors affecting the junior doctor's response on whether to communicate to others are shown below in Figure 1. Escalation is defined here as communication specifically to senior doctors. DCMT starts with the sender's interpretation of the clinical situation and its setting. For instance, a sender may consider a patient-related task request to be routine. The sender's perception of context determines which communication system is used, i.e. the combination of the physical means of communication and mode options. Once a message is received, the recipient then interprets and acts upon it, forming a judgement about clinical risk for the patient, and personal risk for him or herself e.g. criticism by other clinicians.

Communication senders select the ICT system on the basis of their interpretation of usage context, e.g. patient acuity and personal work load. As in Figure 1, the clinical ICT system used then determines the ICT reliability and efficiency of communication receipt. For instance, live video communication may facilitate quick acknowledgement of communication and be faster than describing complex problems. However, it may be less technically reliable than asynchronous text-based systems. Also, recipients may view synchronous communication as disruptive for their own workflow regardless of advantages for senders [23]. ICT reliability also facilitates, but does not obligate, ICT efficiency for both the sender and receiver. An example is that locators may reliably and quickly send protocolised messages for unidirectional emergency group communication to multiple recipients. However, this system may be inefficient in other contexts where bidirectional, rich communication is desirable e.g. evolving situations. Ease of use is an important part of ICT efficiency. Other factors which contribute to ICT efficiency are ICT accessibility (the ability for a user to see the status of other users via the ICT system), and bidirectionality. ICT accessibility and bidirectionality contribute to ICT efficiency although they are features used after or separate to communication receipt, rather than accompanying it. Some features may contribute to ICT efficiency inconsistently e.g. ICT technical design. Whilst increased mobile phone screen size should improve readability, this would be irrelevant if message clarity was inadequate or the message was not received in the first place. Whatever the communication system combination is as in Figure 1, junior doctors ultimately prefer it to deliver maximal ICT reliability, then efficiency. Any other features (e.g. smart mobile phone applications) are secondary considerations.

Communication systems also shape communication content, e.g. locator messages are constrained by character length and type. It is easier to structure and standardise text messaging platforms, which may lead to improved quality of information transfer [34]. Smart mobile phone messages can include photos, improving the efficiency of information transfer. Depending on the situational context, communication content and quality afforded by the communication system may be of low relevance. If a recipient junior doctor had pre-existing awareness of a patient's issues, he or she may require minimal message content to act. Alternatively, doctors may not alter RRS triggering parameters to prevent unnecessary nurse escalation, merely accepting the additional messages [35]. Senders may deliberately vary communication quality according to the intended recipient [12], so that quality is not wholly determined by technological limitations. Junior doctors are seldom free agents in determining which ICT system is used to initiate communication to them and cannot enforce their preferred communication content, shape or timing. They can vary their response, e.g. senders may need to send the same message via multiple communication systems due to lack of recipient acknowledgement.

Whatever the communication content and format, a junior doctor recipient interprets it in terms of risk for both the patient and to themselves as seen in Figure 1. Risk interpretation is influenced by the recipient's responsibility and experience. Responsibility is the recipient's view of their own role, both concerning the perceived patient problem, and the recipient's relationship with senior doctors. Depending on the junior doctor's perception of responsibility, especially to senior doctors, the former may still communicate about a patient's condition despite it being stable. Experience is both the recipient's clinical experience pertinent to the perceived patient issue, and organisational experience. Increased clinical experience would inform recipient knowledge of when to ask for help. This may also be counterproductive in that junior doctors may not communicate because, while they recognise signs of deterioration, they think they can independently manage the deteriorating patient [14], or that they should be able to [36].

Experience arising from previous professional interaction with other clinicians will influence the responsibility felt by the recipient regarding risk communication, e.g. which senior doctors to call and when. When a junior doctor interprets a communication, this is not purely based on their own view of the patient's clinical state, or what theoretical evidence there is [37] – it also depends on problem type definition by both the individual and group [38]. Senior doctors from different specialties may have conflicting views, leaving junior doctors in a difficult intermediary position [38]. These views may be idiosyncratic and defy existing evidence [39]. Consequently, junior doctor experience of their senior doctor expectations is just as important as knowledge of treatment evidence. Experience at clinical communication also improves the efficiency of communicating to other clinicians, e.g. determining who is available and accessible. Recipient risk judgment is what finally determines a decision to initiate communication and to whom it should go to. The desire to communicate efficiently rises as perceived risk does, e.g. communication is then likely to be verbal and to clinicians of direct utility for the patient's care.

The interaction of risk, experience, responsibility and efficiency is complex. An example is that junior doctors with high clinical experience may independently resolve a clinical problem, thereby reducing the likelihood of communicating with senior doctors. Alternately, when clinical and organisational experience are high, routine care doctors may be notified even when not immediately involved. Very junior doctors are less likely to recognise deterioration due to their lack of clinical experience and are not as efficient at patient management or communication. They are also not as knowledgeable about who to communicate to due to lack of organisational experience. Combined with low responsibility, such doctors would be very unlikely to escalate a patient.

The cumulative effect of the context of usage, the clinical ICT system, the communication that ensues and the recipient's interpretation of it is what leads to the latter's decision to initiate further communication. This does not replace the junior doctor's patient assessment and management. Indeed, junior doctors usually do this prior to communicating [11]. DCMT is primarily about the junior doctor's decision to communicate to other clinicians. It is less concerned about what the junior doctor can or should obtain from such interaction, e.g. the merits of activating a group emergency call or calling an adjacent colleague.

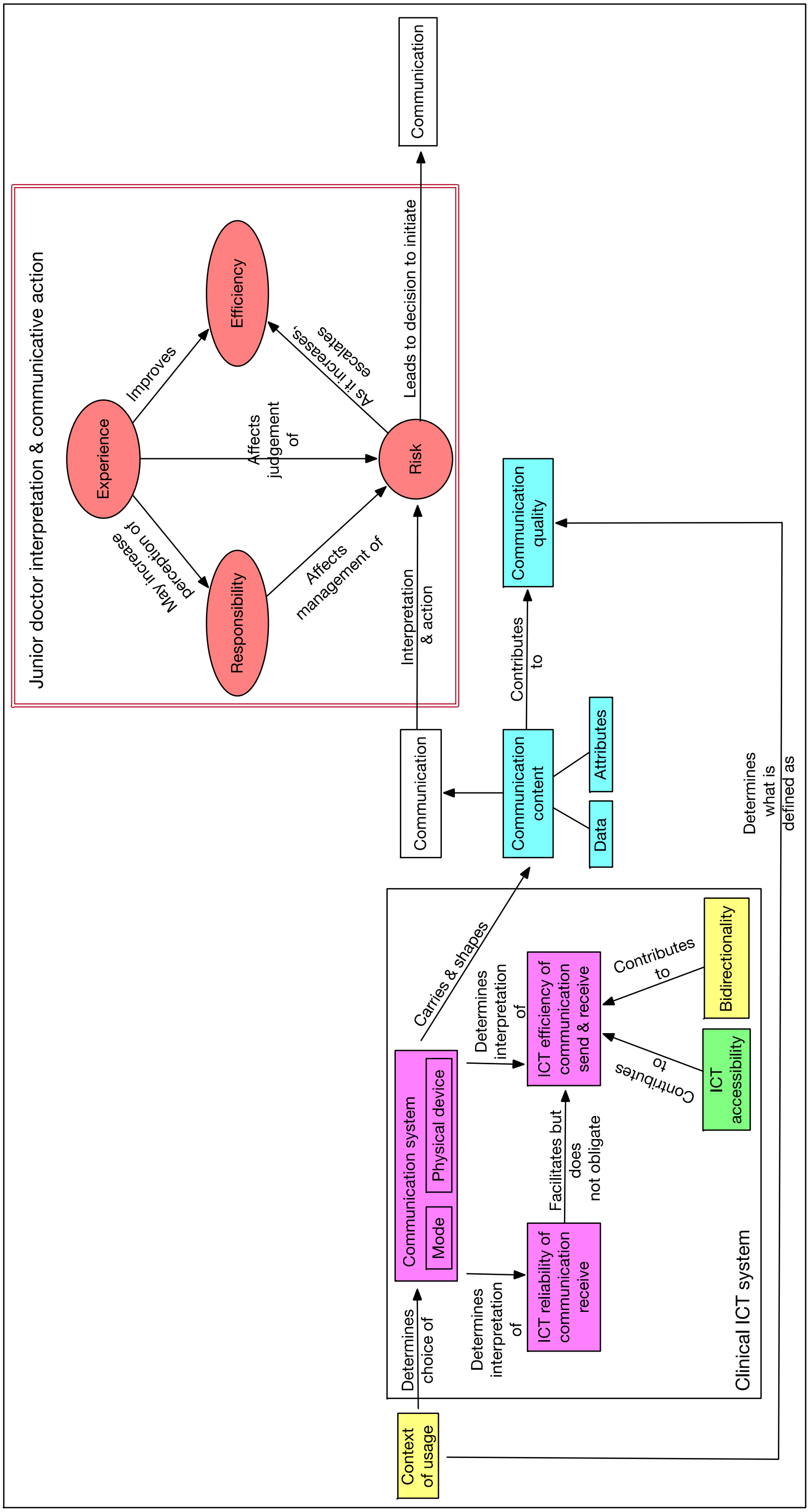


Figure 1. Deterioration Communication Management Theory

3. Usage of DCMT in relation to present theories regarding junior doctor communication

Grounded theory (GT) and action research (AR) were the methodologies used in the research behind DCMT [18]. There are general advantages to using GT in healthcare organisation research. In addition, there were specific reasons why GT was used in combination with AR to develop DCMT. These issues and a discussion of GT version selection in DCMT research are below.

GT is commonly used in healthcare research. Its advantages include its innate pragmatism, adaptability to study resources and where there are few pre-existing theories [40]. This is especially beneficial in complex organisations with unpredictable phenomena. Other researchers of junior doctor communication for the purposes of requesting assistance [11, 12] have also used GT. GT can also be used together with action research (AR), such that the two methodologies complement each other [41]. GT enhances the rigour of AR theory building and chain of evidence creation [42], whilst AR helps address power differences between researcher and researched [18]. GT was also particularly useful in DCMT research in that theory generated should be modifiable as new data arises. This is so that its concepts are modified as necessary and the theory continues to be relevant [43]. This is particularly important as hospital organisations and ICT systems evolve with time.

Classical GT (CGT) is one of three common versions of GT, the other two being Straussian GT and constructivist GT. The versions differ in their underlying philosophy, coding techniques and approach to research of previous literature [44]. An example of GT version selection according to research question is Tallentire et al's [11] usage of constructivist GT to understand junior doctor behaviour in acute patient care. Liang [18] used CGT to explain junior doctor communicative behaviour in response to messages. This is consistent with differences between these GT versions, in that constructivist GT aims to understand multiple perspectives in a social process, whilst CGT focuses on one main concern [43]. CGT offers other advantages in both healthcare [43] and information systems [40] research. CGT is flexible in accommodating many types and sources of data. This is advantageous in hospital ICT research given the variety of ICT physical devices and modes used [22]. CGT can accommodate different theoretical perspectives. Patient safety research, of which DCMT is an example, has classically favoured positivistic and quantitative approaches. These do not reflect complexity or explain underlying factors well [45]. Whilst CGT as a GT version has seldom been used in information systems research [46], its philosophical pragmatism is advantageous in bridging these two worlds. CGT generated theories also have theoretical concepts transferrable to different settings [47], meaning that DCMT is potentially applicable to other organisational situations where individuals communicate about risk.

4. Explanation of success or failure in communication about the deteriorating patient

Since DCMT has not been used in ICT development as yet, its potential usage will be demonstrated with an example of a junior doctor ICT system. Patel et al [34] studied information transfer and clinician satisfaction with locators compared to Hark, a mobile phone-based task management system. Nurses sent pre-arranged messages or locator pages to junior doctors during usual hospital business days. Recipient reply was defined as either reading the Hark message or calling the locator page number. Further communication from the junior doctors to any other clinician was not studied. Participant communication was assessed by the Quality of Information Transfer (QUIT) tool and participation satisfaction was assessed by survey. Message receipt speed was no different between Hark and locators, although the former was thought to have better information transfer quality and overall ICT system efficiency. Participants did not prefer Hark to locators as a primary system for ward nurse to doctor communication. This is surprising, given junior doctors often prefer mobile phones [23].

DCMT provides another lens by which to interpret and explain Patel et al's [34] findings, via ICT efficiency as in Figure 1. Hark and locators had similar ICT reliability in that both sent messages equally fast, although Hark was thought to have fewer errors. Hark facilitated ICT efficiency in terms of prioritisation features, information storage and transfer, acting as a communication record, facilitating collaboration and minimising interruption. However, Hark did not offer improvement in other aspects of ICT efficiency. The two ICT systems were equivalent with regard to sufficient communication detail, ease of use and the ease of contacting other clinicians. Hark messages did not indicate which senior doctor was responsible for the patients care and thus the recipient junior doctor would not have known which other team doctors to communicate to. Findings equivalent to two other DCMT factors contributing to ICT efficiency were also equivocal. The exact nature of ICT accessibility was not stated. Although junior doctors might first receive urgent messages asynchronously, they usually prefer synchronous communication for urgent issues [23]. It is not clear which ICT system mode would be used, i.e. whether Hark mobiles permitted telephony [34] and thus its accessibility. Hark did allow for a reply to the original question but did not appear to facilitate discussion beyond this, i.e. it lacked full ICT bi-directionality. The cumulative effect would be that although Hark was partially better than locators, it was not clearly superior in ICT reliability or efficiency. This may be why participants did not prefer Hark as a primary ICT system [34], and suggests how Hark could be improved.

Kelly et al's [36] paper will be used to demonstrate how DCMT can explain junior doctor interpretation and response to communication about the deteriorating patient, as Patel et al [34] focused on ICT evaluation. Kelly et al [36] surveyed junior doctors in one hospital regarding barriers and enablers to deteriorating patient escalation. The major barriers would be classified under DCMT into responsibility identification, experience at communication, assessment and patient management, and risk. Risk was that associated with determining which patients needed to be escalated and the response of senior doctors. Kelly et al.'s [36] participants also thought that risk judgment could be impaired by contextual factors e.g. competing demands. In DCMT, risk is an interpretation by junior doctors receiving information rather than an absolute and objective representation of a patient's risk of deterioration. Consequently, improving adverse contextual factors may not decrease escalation. Concurrent high workload, by decreasing time available for patient care, may increase perceived patient risk and thus favour escalation. Kelly et al.'s participants [36] suggested countering communication barriers by improving junior doctor training and organisational culture e.g. senior doctor accessibility. ICT improvement was also suggested as a minor factor. In practice, these changes may be complex to achieve [11], let alone to any consistent degree between different hospitals [36].

DCMT analysis of Kelly et al.'s [36] major escalation barriers suggests other targets for improvement. Clarification of communication participants and task responsibility is important for all health ICT systems and especially so for the treatment of deteriorating patients [23]. ICT systems can augment and control group awareness of patient deterioration, e.g. by preferentially directing alerts to more experienced clinicians. What DCMT shows is that as far as junior doctors are concerned, such ICT systems need to be designed with thought to sociotechnical as well as clinical issues to bring about the desired effects. In addition, no present deteriorating patient risk indicator system is clearly the best in detection and prediction of clinical outcome. Present RRS risk indicators for communication are primarily based on physiological indices or subjective concern and lack accuracy [1]. ICT systems can be used to improve accuracy and presentation, but not without first addressing their deficiencies in ICT efficiency [23].

5. Conclusion

DCMT is a new theory which can be applied to improve existing ICT systems used for ward communication and to configure them for the deteriorating patient context. It has not previously been applied. As the research on which it was based was mostly done at one hospital, usage would help to confirm or improve its transferability. Its treatment of ICT system selection as a combination of physical device, software and context is useful in that each can change over the course of a patient's deterioration. Although DCMT can be used in the design of a standalone ICT system, hospitals and individuals usually have a combination of ICT systems. Consequently, DCMT is better applied to optimise communication originating from the sender's interpretation of clinical context to the entirety of ICT systems which could be used, rather than only considering the design of one.

DCMT also addresses what occurs when a junior doctor receives communication about a deteriorating patient, and whether they decide to communicate in turn. It is affected by social and organisational factors and is not solely dependent on an individual junior doctor's knowledge. While DCMT does not specify how to improve these factors, it can be used to predict their effect on junior doctor communication and thereby to improve the chances of appropriate escalation of deteriorating patients.

Teaching questions for reflection

1. Would highly accurate prediction of patient risk affect junior doctor escalation of patients?
2. Is it possible, or desirable, to standardise communication content for the purpose of improving risk interpretation?
3. If ICT systems automate group awareness of patient deterioration, will junior doctors independently escalate patients?

References

[1] K. White, I.A. Scott, A.Vaux, C.M. Sullivan, Rapid response teams in adult hospitals: time for another look ? *Intern Med J*, **45** (2015), 1211-1220.

[2] D. Massey, L.M. Aitken, W. Chaboyer, What factors influence suboptimal ward care in the acutely ill ward patient ?, *Intensive Crit Care Nurs*, **25** (2009), 169-180.

[3] S. Romero-Brufau, J.M. Huddleston, G.J. Escobar, M. Liebow, Why the C-statistic is not informative to evaluate early warning scores and what metrics to use, *Crit Care,* **19** (2015), 285-290.

[4] M.N. Storm-Versloot, L. Verweij, C. Lucas, J. Ludikhuize, J.C. Goslings, D.A. Legemate, Clinical relevance of routinely measured vital signs in hospitalized patients: a systematic review, *J Nurs Scholarsh*, **46** (2014), 39-49.
[5] N.E. Pedersen, L.S. Rasmussen, J.A. Petersem, T.A. Gerds, D. Ostergaard, A. Lippert, A critical assessment of early warning score records in 168,000 patients, *J Clin Monit Comput*, **81** (2017), 109-116.
[6] A.L. Benin, C.P. Borgstrom, G.Y. Jenq, S.A.Roumanis, L.I. Horwitz, Defining impact of a rapid response team: qualitative study with nurses, physicians and hospital administrators, *BMJ Qual Saf*, **21** (2012), 391-398.
[7] Health Information and Quality Authority, Health technology assessment of the use of information technology for early warning and clinical handover systems, Health Information and Quality Authority Dublin, 2015.
[8] A. Callaghan, L. Kinsman, S. Cooper, N. Radomski, The factors that influence junior doctors' capacity to recognize, respond and manage patient deterioration in an acute ward setting: an integrative review, *Aust Crit Care*, **30** (2016), 197-209.
[9] C.M. Smith, G.D. Perkins, I. Bullock, J.F. Bion, Undergraduate training in the care of the acutely ill patient: a literature review, *Intensive Care Med*, **33** (2007), 901-907.
[10] L.V. Monrouxe, A. Bullock, G. Gormley, K. Kaufhold, N. Kelly, C.E. Roberts et al, New graduate doctors' preparedness for practice: a multistakeholder, multicentre narrative study, *BMJ Open*, **8** (2018), e023146.
[11] V.R. Tallentire, S.E. Smith, J. Skinner, H.S. Cameron, Understanding the behaviour of newly qualified doctors in acute care contexts, *Med Educ* **45** (2011), 995-1005.
[12] J. Martland, D. Chamberlain, A. Hutton, M. Smigielski, Communication and general concern criterion prior to activation of the rapid response team: a grounded theory, *Aust Health Rev* **40** (2016), 477-483.
[13] N. Credland, J. Dyson, M.J. Johnson, What are the patterns of compliance with early warning track and trigger tools: a narrative review, *Appl Nurs Res*, **44** (2018), 39-47.
[14] D. Palakshappa, L.P. Carter, C.M. El Saleeby, Discrepancies in after-hours communication attitudes between pediatric residents and supervising physicians, *J Pediatr*, **167** (2015), 1429-1435.e2.
[15] C. O'Donnell, S. Thomas, C. Johnson, L. Verma, J. Bae, D. Gallager, Incorporating patient acuity rating score into patient handoffs and the correlation with rapid responses and unexpected ICU transfers, *Am J Med Qual*, **32** (2017), 122-128.
[16] V.M. Arora, J.K. Johnson, D.O. Meltzer, H.J. Humphrey, A theoretical framework and competency-based approach to improving handoffs, *Qual Saf Health Care*, **17** (2008), 11-14.
[17] J. Abraham, T. Kannampallil, V.L. Patel, A systematic review of the literature on the evaluation of handoff tools: implications for research and practice, *JAMIA*, **21** (2014), 154-162.
[18] J. Liang, *How technology can improve junior doctor communication about the deteriorating patient*, University of Auckland, Auckland, 2018.
[19] S. Quan, R.C. Wu, P.G. Rossos, T. Arany, S. Groe, D. Morra et al, It's not about pager replacement: an in-depth look at the interprofessional nature of communication in healthcare, *J Hosp Med,* **8** (2013), 137-143.
[20] J.M. Weller, M. Barrow, S. Gasquoine, Interprofessional collaboration among junior doctors and nurses in the hospital setting, *Med Educ*, **45** (2011), 478-487.
[21] J. Hill, M. Janko, K.F. Angell, K. Day, M.G. Noujaim, J. Kalil et al, Hospital text paging communication as a surgical quality improvement initiative, *J Surg Res*, **213** (2017), 84-89.
[22] R.C. Wu, K. Tran, V. Lo, K.J. O'Leary, D. Morra, S.D. Quan, L. Perrier, Effects of clinical communication interventions in hospitals: a systematic review of information and communication technology adoptions for improved communication between clinicians, *Int J Med Inform,* **81** (2012), 723-732.
[23] R.C. Wu, V. Lo, D. Morra, B.M. Wong, R. Sargeant, K. Locke, et al, The intended and unintended consequences of communication systems on general internal medicine inpatient care delivery: a prospective observational case study of five teaching hospitals, *JAMIA*, **20** (2013), 766-777.
[24] I. Popovici, P.P. Morita, D. Doran, S. Lapinsky, D. Morra, A. Shier, et al, Technological aspects of hospital communication challenges: an observational study, *Int J Qual Health Care*, **27** (2015), 183-188.
[25] D.G. Morrow, U.M. Fischer, Communication in socio-technical systems, in J.D. Lee, A. Kirlik, *The Oxford Handbook of Cognitive Engineering*, Oxford University Press, New York, 2013.
[26] C. Boissin, L. Blom, L. Wallis, L. Laflamme, Image-based teleconsultation using smartphones or tablets: qualitative assessment of medical experts, *Emerg Med J*, **34** (2017), 95-99.
[27] N. Kapur, A. Parand, T. Soukup, T. Reader, N. Sevdalis, Aviation and healthcare: a comparative review with implications for patient safety, *J R Soc Med*, **7** (2016), 1-10.
[28] D.G. Morrow, K.D. Lopez, Chapter 3: Theoretical foundations for health communication research and practice, *Cognitive Informatics for Biomedicine*, Springer International Publishing, Switzerland, 2015.

[29] M.J. Johnston, D. King, S. Arora, K. Cooper, N.A. Panda, R. Gosling et al, Requirements of a new communication technology for handover and the escalation of patient care: a multi-stakeholder analysis, *J Eval Clin Pract,* **20** (2014), 486-497.

[30] S. Rokadiya, J.A. McCaul, D.A. Mitchell, P.A. Brennan, Use of smartphones to pass on information about patients – what are the current issues, *Br J Oral Maxillofac Surg*, **54** (2016), 596-599.

[31] J. Stewart, To call or not to call: a judgement of risk by pre-registration house officers, *Med Educ*, **42** (2008), 938-944.

[32] R. Janss, S. Rispens, M. Segers, K.A. Jehn, What is happening under the surface ? Power, conflict and the performance of medical teams, *Med Educ*, **46** (2012), 838-849.

[33] B.G. Glaser, Choosing grounded theory, *The Grounded Theory Review*, **13** (2014), 3-19.

[34] B. Patel, M. Johnston, N. Cookson, D. King, S. Arora, A. Darzi, Interprofessional communication of clinicians using a mobile phone app: a randomized crossover trial using simulated patients, *J Med Internet Res,* **18** (2016), e79.

[35] C. Douglas, S. Osbourne, C. Windsor, R. Fox, C. Booker, L. Jones et al, Nursing and medical perceptions of a hospital rapid response system: new process but same old game ?, *J Nurs Care Qual*, **31** (2016), E1-E10.

[36] C. Kelly, S. Larwill, L. Hamley, M. Sandford, Failure to escalate: what stops junior doctors asking for help when they need it ?, *APJHM,* **9** (2014), 38-46.

[37] M.S. Nilsson, E. Pilhammar, Professional approaches in clinical judgements among senior and junior doctors: implications for medical education, *BMC Med Educ*, **9** (2009), 25.

[38] S. Bull, K. Mattick, K. Postlethwaite, 'Junior doctor decision making: isn't that an oxymoron ?' A qualitative analysis of junior doctors' ward-based decision-making, *J Vocat Educ Train*, **65** (2013), 402-421.

[39] E. Charani, E. Castro-Sanchez, N. Sevdalis, Y. Kyratsis, L. Drumright, N. Shah et al, Understanding the determinants of antimicrobial prescribing within hospitals: the role of "prescribing etiquette", *Clin Infect Dis*, **57** (2013), 188-196.

[40] D.F. Birks, W. Fernandez, N. Levina, S. Nasirin, Grounded theory method in information systems research: its nature, diversity and opportunities, *Eur J Inf Syst,* **22** (2013), 1-8.

[41] B. Dick, What can grounded theorists and action researchers learn from each other, in The SAGE Handbook of Grounded Theory, SAGE England, (2007).

[42] R. Baskerville, J Pries-Heje, Grounded action research: a method for understanding IT in practice, *Accting, Mgmt & Info Tech*, **9** (1999), 1-23.

[43] J. Breckenridge, D. Jones, I. Elliot, M. Nichol, Choosing a methodological path: reflections on the constructivist turn, *The Grounded Theory Review*, **11** (2012), 64-71.

[44] M. Kenny, R. Fourie, Contrasting Classic, Straussian, and Constructivist Grounded Theory: methodological and philosophical conflicts, *Qual Rep,* **20** (2015), 1270-1289.

[45] W.B. Runciman, G.R. Baker, P. Michel, I.L. Jauregui, R.J. Lilford, A. Andermann, R. Flin, W.B. Weeks, The epistemology of patient safety research, *Int J Evid Based Healthc,* **6** (2008): 476-486.

[46] R. Matavire, I. Brown, Investigating the use of "Grounded Theory" in information systems research, *Proceedings of the 2008 annual research conference of the South African Institute of Computer Scientists and Information Technologists on IT research in developing countries riding the wave of technology - SAICSIT '08*, New York, 139-147.

[47] B.G. Glaser, Conceptualization: on theory and theorizing using grounded theory, *Int J Qual Methods*, **1** (2002), 23-38.

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

doi:10.3233/SHTI190118

Resilient Health Care: A Determinant Framework for Understanding Variation in Everyday Work and Designing Sustainable Digital Health Systems

Robyn CLAY-WILLIAMS[a,1] and Jeffrey BRAITHWAITE[a]
[a] *Macquarie University, Sydney Australia*

Abstract. This chapter presents an overview of Resilient Health Care (RHC), introducing two aspects of RHC that are important for designing sustainable digital health systems and for considering implementation outcomes: (1) understanding how normal variation in everyday work can affect implementation of digital health interventions, and (2) the role of information systems in coping with unexpected events. The importance of considering how variation in everyday work can lead to wanted and unwanted outcomes when designing information systems is illustrated through a case study of implementation of a telehealth intervention. We examine how normal variation in everyday work can lead to both safety and error, and discuss how consideration of system resilience when designing and implementing health informatics applications can contribute to improving safety for patients in the future. How health information systems can assist organisations in coping with the unexpected is illustrated through a second case study, of a thunderstorm asthma event in Melbourne, Australia. We briefly present the thunderstorm asthma case, and discuss the role of healthcare informatics in preparing for future unexpected events affecting population health.

Keywords. Resilient Health Care, Patient Safety, Complex Adaptive System, Safety-I, Safety-II

Learning objectives

After reading this chapter the reader will:

1. Understand the background to Resilient Health Care (RHC) and its historical antecedents.
2. Appreciate the main currents and selected underlying concepts in the field, including Safety-I and Safety-II; and Work-as-Imagined and Work-as-Done.
3. Apply knowledge about RHC to current research-based or practice-based problems in health informatics.
4. Analyse health informatics problems in a frame that offers a more positive vision of how safe, effective care can be delivered in complex, dynamic health settings.
5. Consider normal variation in everyday work when designing or implementing health informatics systems.

[1] Corresponding Author, Robyn Clay-Williams, E-mail: robyn.clay-williams@mq.edu.au

1. The scope of Resilient Health Care

When designing and implementing new digital health systems, the safety of those systems for clinicians and the patients in their care must be a core consideration. Resilient Health Care (RHC) is a relatively new approach to safety, albeit with long antecedents to resilience engineering, that shifts from understanding safety as the absence of accidents or incidents, to thinking of safety as a system where as many things as possible go well. Measuring what goes wrong has been an attractive concept for organisations in the past: there are typically few things to count, and resources can usually be brought to bear to tackle problems that have been shown to result in significant harm. Traditional approaches to safety are reactive rather than proactive; examples include regulation,[1] protocols and checklists,[2-5] Root Cause Analyses[6] and judicial inquiries to investigate patient deaths.[7] Errors after they occur are identified and rectified, and processes are put in place in an attempt to prevent future occurrences. This approach is not effective, however, where the route to error is different on each occasion, and where fixes for previous errors can contribute to new paths to failure. In contrast, RHC asks us to understand how the systems requiring action actually work, to identify what goes right and comprehend why things routinely go well, and to proactively manage variability in the workplace. This newer way of thinking is necessary for the whole gamut of systems behaviour. It is especially apposite for improving safety in complex adaptive systems such as healthcare, and has been driven by failure to improve the safety of patient care by traditional means, despite more than two decades of effort.[8]

A complex adaptive system is one with multiple interacting and interdependent parts that change continuously and dynamically in response to environment or conditions.[9] In healthcare, these components consist primarily of humans, such as clinicians, patients and their families, aided by affordances such as technological artefacts and equipment. Human performance is inherently variable; regardless of their experience and ability, for example, the performance of an individual clinician will vary depending on the problem, time of day, and so on. Furthermore, clinicians work in small, medium and large ad hoc teams, and must interact with a range of other healthcare professionals whose performance is also varying. When the variability associated with patients and their illness or injury is also taken into account, the result is a complex and unpredictable system. Due to the complex and dynamic nature of the interactions of components, outcomes from a complex adaptive system can be unexpected and unable to be attributed to specific inputs—this is what is known as *emergent* behaviour. In addition, the system's history plays a part in determining where things are now; this is called 'path dependence'.[10]

As RHC grew from the field of resilience engineering, it borrowed from resilience engineering theory, which conceptualises how normal variation in task performance in socio-technical systems can lead to both wanted and unwanted outcomes. This guides research into how variation in human performance of everyday work processes contributes to both failure (i.e. unwanted outcome or 'error') and success (i.e. wanted outcome). The theory is grounded in system thinking and complexity science, and in understanding how systems typically cope successfully with unwanted outcomes (or events) that are unexpected. Resilience engineering originated in 2005, at a gathering of influential industrial safety scholars led by Erik Hollnagel, David Woods and Nancy Leveson,[10] and emerged from the work of Crawford Holling on ecological systems[11] and Charles Perrow on normal accidents.[12] The application of resilience engineering principles to healthcare can be traced to a meeting in 2012 of resilience engineering and healthcare safety experts led by Hollnagel, Jeffrey Braithwaite and Robert Wears and has since grown to involve a large and increasingly influential group, the Resilience Health Care Network (https://resilienthealthcare.net).[13] In the field of RHC, resilience is defined as the ability of the health care system (a clinic, a ward, a hospital, a country) to adjust its functioning prior to, during, or following events (changes, disturbances, and opportunities), and thereby sustain required operations under both expected and unexpected conditions.[14]

RHC is identified with two complementary approaches to safety – Safety-I and Safety-II. Neither approach is superior, however one approach might work better than the other depending on the complexity and predictability of the situation. Safety-I is an approach that is effective for minimising error in linear systems, where the interaction between components is well characterised, resulting in well-defined and predictable outcomes. Linear systems can range from simple to complicated, but the system outcome can always be predicted with a high degree of certainty provided we know the system inputs. In linear systems, the boundaries are usually fixed or able to be clearly defined, which means that local problems can be addressed independently of the larger system, and solutions can be generalised.

The best examples of linear systems are systems with primarily technological components, such as the computerised aspects of a digital heath system, or an anaesthetic machine. For an anaesthetic machine we understand how each of the electronic and mechanical parts are connected and operate so that the machine can function, and we can often predict accurately the mean time between failure for these sub-components. For a linear system, process-oriented controls such as standardisation of manufacture and operation provide effective safety measures, and barriers to error propagation across such a system can be applied effectively.

Once we add a sociological component, such as normal human behaviour, into the system, it becomes more complex, and Safety-I solutions become less effective. In contrast, Safety-II is an approach that is suited to a complex system. Rather than focusing on failures, Safety-II thinking tries to understand how human performance nearly always goes well and leverages that information to improve the number of things that go right. In a complex system, boundaries can be porous, and there is significant interaction between local context and the larger system. Rather than adding system controls or barriers, which is difficult to do when boundaries are not well-defined, a Safety-II approach will try to simplify the system and rely on the adaptability of the humans in the system to adjust their performance in response to changing system demands.

To apply RHC principles in the workplace to improve the number of things that go right, we need to understand 'Work-as-Done', or how clinicians make continuous small and large adjustments during their daily work, to satisfy the changing needs of patient care. In complex systems, 'Work-as-Done' is usually different to Work-as-Imagined' by those who administer healthcare and who develop the rules and procedures that clinicians must follow. This can result in different assumptions across hospitals of how tasks are accomplished, and can make implementation of new processes and procedures difficult and, sometimes, unsafe for patients. A digital health system that is designed without in-depth knowledge of how everyday work is accomplished may not be usable by clinicians, and result in clinician frustration and workarounds.

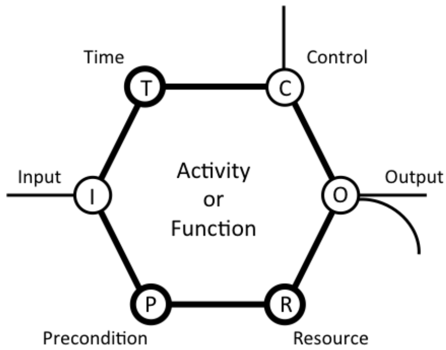
In terms of implementation science, RHC can be considered a *determinant framework*[15] that helps us to design and implement successful interventions through understanding healthcare professionals and the system in which they work. The tools of RHC, while still in early stages of development, have potential to complement other determinant frameworks such as computational simulation modelling (e.g. system dynamic modeling,[16-18] discrete event modelling[19, 20] and agent based modelling[21, 22]). This chapter presents two aspects of RHC applicable to interventions in health informatics: understanding how normal variation in everyday work can affect design and implementation of sustainable digital health systems, and designing information systems to cope with unexpected events.

2. Applications of Resilient Health Care in health informatics

2.1. Identifying and understanding variability in everyday work

The importance of considering wanted and unwanted variation in everyday work when designing sustainable digital health systems is illustrated through a case study of the implementation of an Australia-wide video consultation and triage service supporting expecting parents and parents, families and carers of young children. Established in 2010, the telehealth service consists of a national helpline, video and website service sponsored by the Australian government. Telephone consultation and triage services are commonly used to deliver health advice worldwide. In Australia, availability of high-speed internet services in remote areas is driving a move from telephone to video telehealth services for healthcare providers; however, providers are unfamiliar with how to introduce and operate a video service. When designing a new system of work, it is important to take into consideration how day-to-day work is currently carried out, in order to improve uptake and reduce workarounds when the system is implemented.[23] A useful tool for understanding variation in everyday work, including how that variation in combination with multiple interacting activities can affect outcomes, is the Functional Resonance Analysis Method (FRAM).[24]

The FRAM supports modelling complex socio-technical systems and is developed by determining the activities or functions that make up a process, and how they are coupled. Depending on the problem to be solved or question to be answered, the process can be modelled broadly, or at a more detailed level. For example, if we wanted to model the processes involved in using an automatic teller machine (ATM), we might break the process broadly into activities of (1) insert card, (2) enter PIN, (3) enter withdrawal amount, and (4) take money and card. However, if we were interested in specific detail such as the usability of the ATM screen, we might expand step (3) to include additional steps for select savings account, check account balance, enter withdrawal amount, request receipt, and so on. The data for developing a FRAM model can be obtained through a number of methods, including ethnography, interviews, documented processes, and so on. Each function is then described in terms of six aspects (see Figure 1):



Input (I) is what the function acts on or changes
Output (O) is what emerges from the function
Precondition (P) is what must be satisfied before the function can begin
Resources (R) are materials/people needed to carry out the function
Control (C) is how the function is monitored or controlled
Time (T) is any time constraints that might affect the function

Figure 1. FRAM activity hexagon [25] .

A FRAM model is built using a software tool called the FRAM Model Visualiser (FMV).[25] The potential variability of each activity is annotated as the model is built, and can be defined in terms of source of variability (internal or external, type, likelihood), output with regard to time (too early, on time, too late, not at all), and output with regard to precision (possible but unlikely, typical, possible and likely). The resulting model can be interpreted to determine how variability present in each activity affects other activities, and how delays can propagate through the system. Such a model can help to predict unwanted variation when the new system is implemented.

In the telehealth service case, two levels of direct client support are provided: (1) Customer Support Officers (CSOs) provide standardised advice on common situations, such as planning for pregnancy, foods to avoid when pregnant, and breastfeeding, and (2) accredited counsellors provide psychological support and counselling. To illustrate where we found variability in Work-as-Done in our telehealth evaluation, Figure 2 is a FRAM model showing the portion of the work activity where calls are answered and dispositioned by the CSO (we have simplified the FRAM for ease of interpretation, and have not included Resource, Control or Time aspects in the figure). Calls answered and resolved by the CSO form a linear process, passing through steps 1 to 4 (shadowed steps). Variation is indicated in the model by the sine curve within the function (see steps 2 and 4). In this case, the time taken to chat with the client to establish the purpose of the call (2) can vary depending on the client, the purpose of the call, and the expertise of the CSO. The time taken to resolve the call (4) can also vary depending on the complexity of the problem raised by the client, and the amount of information that must be passed from CSO to client to resolve the issue. Once the CSO has resolved the problem (Step 4), they are then available to return to Step 1 to take the next call (CSO availability shown as a precondition).

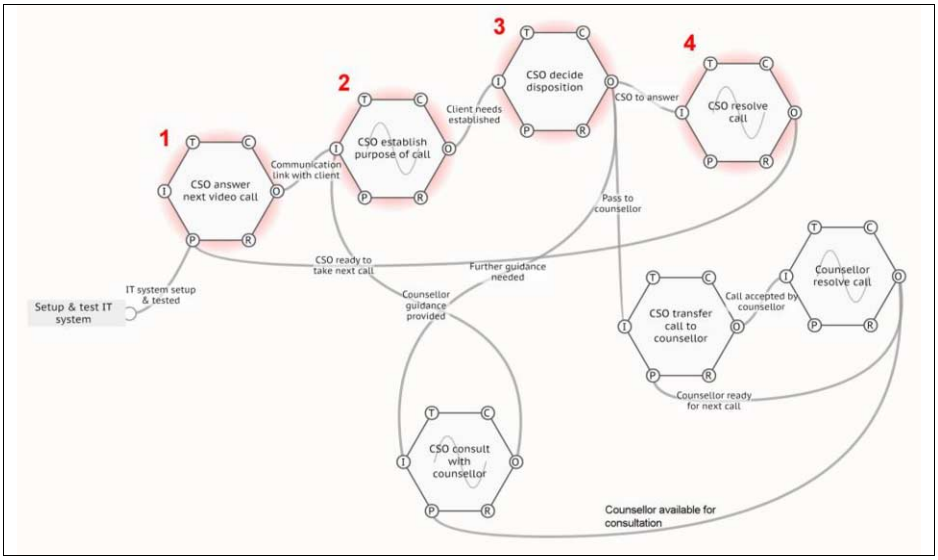


Figure 2. Resolving calls – CSO.

Figure 3 is the same FRAM model showing the portion of the work activity where calls are passed to the counsellor for resolution. It is easy to see from the Figure that including only one additional person in the process increases the complexity and resulting variation. In this process, the CSO takes the call from the client (1), establishes the purpose of the call (2) and decides whether it needs to be passed to a counsellor (3). Sometimes the CSO lacks sufficient expertise, or is uncertain about the correct disposition, so must consult with a counsellor (4) to obtain more information (5) and make the decision (6). The call can then be passed to the counsellor (7), who will resolve the client issue (8). Variation is evident in terms of the time for the CSO to establish the problem (2), in consultation with the counsellor if necessary (3, 4, 5); and for the counsellor to resolve the call with the client (8). We will also see interactions between functions that can exacerbate variability: for example, the counsellor must be available to give advice at step 4, and not on another call. Otherwise the CSO must either wait, or seek advice from another counsellor. We can see where workarounds might arise: if, for example, all counsellors are on other calls, the CSO may decide to proceed without advice, potentially leading to incorrect disposition. We can also see how the advice loop (steps 3-4-5-6) could consume CSO and counsellor time, leading to delays in providing advice by counsellors (8), and backup of new calls waiting for the CSO (1, precondition).

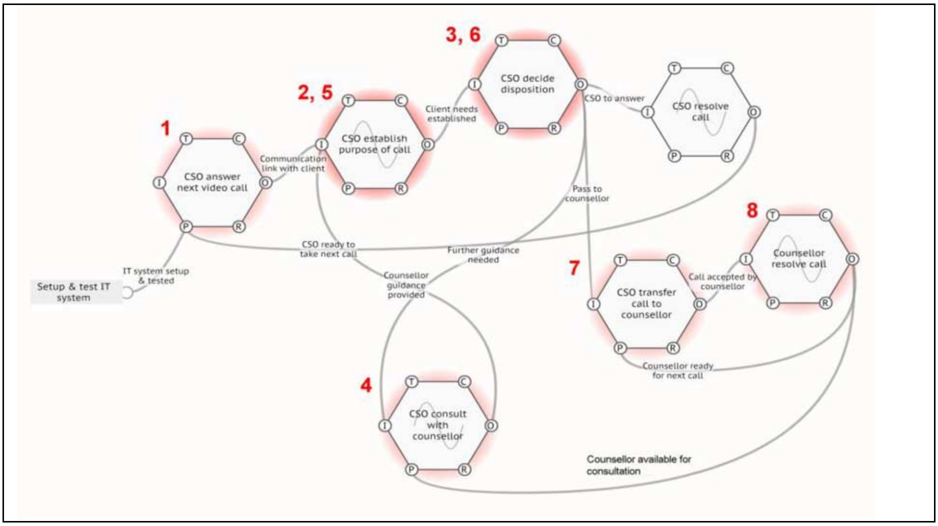


Figure 3. Resolving calls – counsellor.

2.2. Designing information systems to cope with unexpected events

How health informatics can enable systems to cope with the unexpected will be illustrated through a case study of a thunderstorm asthma event in Melbourne, Australia.[26] Over two days in November 2016, nearly 10,000 people presented at hospital Emergency Departments with breathing difficulties, and nine people died. The efficiency and effectiveness of locally embedded health information networks enabled emergency services to manage the unanticipated increase in ambulance calls and hospital presentations, however the crisis revealed deficiencies in command and control level information systems. A useful tool for proactive evaluation of resilience in response to unexpected events is the Resilience Assessment Grid (RAG).[27]

The RAG was derived by considering four essential capabilities of resilience (Figure 4): knowing what to do in response to unexpected occurrences and being capable of doing it (actual), knowing how to identify early that developing events might prove problematic (critical), knowing what to expect as events develop (potential), and learning from what has happened in the past (factual). The ability to respond includes taking unpredictability into account and adjusting responses to enable local experts to improvise. The ability to monitor includes tracking how things are being done well and understanding Work-as-Done. The ability to anticipate includes policy makers balancing prescriptive controls with local level discretion, improvisation and judgement. The ability to learn should be based on frequency and severity of what goes right.

The RAG can be proactively applied by evaluating an organisation in terms of the four capabilities. This evaluation is usually completed as a series of probing questions that can be answered via a combination of interviews, focus groups, ethnography and audit or document review.

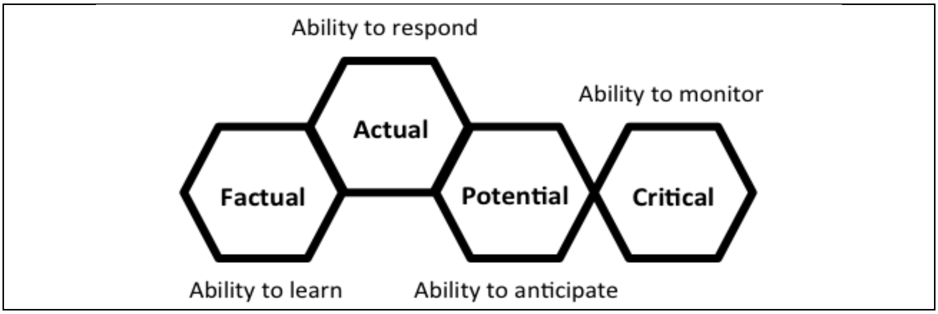


Figure 4. The four resilience potentials forming the RAG [27].

In terms of the thunderstorm asthma case, and guided by the RAG, we can examine what happened retrospectively in order to learn and develop information systems to improve the response of emergency services to future large scale unexpected events affecting population health. One of the difficulties faced by emergency services was that, at the time people were experiencing acute respiratory symptoms, information on the cause and extent of the problem was limited. Using the RAG framework, we can look at the successes and failures of information systems when challenged by the thunderstorm asthma event, as follows:

Actual: the ability to respond to the thunderstorm asthma event. Despite the rapid onset of events, the Emergency Services Telecommunications Authority (ESTA), Ambulance Victoria (AV) and Victorian hospitals responded quickly and increased the scale of their respective operations. The State Health Emergency Response Plan (SHERP) was not activated at an appropriate level, however, so processes to aggregate and share data were not available. In addition, neither ESTA nor AV formally activated their emergency escalation plans. The key decision-maker was the State Health and Medical Commander (DHHS). DHHS communicated with hospitals through mobile text messages, phone calls and emails to individuals such as hospital Chief Executive Officers; this resulted in inefficiency, and inconsistency of information provided to hospitals. In response, some hospitals contacted each other directly to obtain information.

Critical: the ability to monitor as thunderstorm asthma developed. When information is limited, it is vital to identify triggers for action. During the thunderstorm asthma event, there was a surge in demand for telecommunications, ambulance and hospital services. Monitoring of usage by the Emergency Services Telecommunications Authority (ESTA; see Figure 5), Ambulance Victoria (AV) and Victorian hospitals for future unexpected events may allow for a rapid surge in demand to act as a trigger to activate emergency response plans.

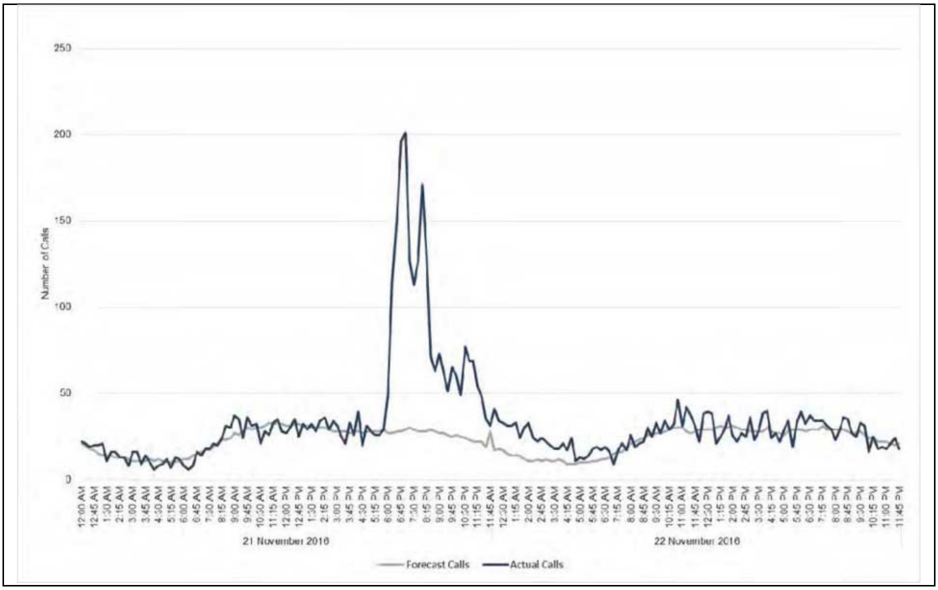


Figure 5. Emergency calls for ambulances presented to ESTA on 21-22 November [26].

Potential: the ability to anticipate the severity of the asthma crisis. Despite a forecast for severe thunderstorms, there was no expectation of an impending emergency. Lack of situational awareness across the health system meant that, although clinicians suspected that the respiratory symptoms they were seeing were caused by thunderstorm asthma, there was no channel for sharing this information with DHHS. In addition, the traditional system for communicating public health concerns, whereby DHHS seeks to understand what is causing the problem in combination with its impact on the health system before issuing public information and warnings, was unsuited to a rapid-onset problem such as thunderstorm asthma.

Factual: the ability to learn from successes and mistakes when responding to the thunderstorm asthma event. Following the event, the state Inspector-General for Emergency Management was tasked by the state government to review the emergency response.[26] The review resulted in 16 comprehensive recommendations, of which 10 were related to improving data integration and/or information systems. Various organisations that were part of the response, including ESTA and AV, also reviewed and updated their emergency response plans. Finally, an interagency working group was established to share knowledge and improve procedures for detecting and anticipating the severity of future events.

3. Explanation of success or failure of health IT system

3.1. Identifying and understanding variability in everyday work

Using RHC principles enabled an understanding of Work-as-Done when delivering telehealth advice to new and prospective parents via video. The plan to use both CSOs and counsellors to deliver the service was abandoned, and a revised system design whereby CSOs continued to deliver their service over the telephone and only counsellors participated in the video service. This decision was made prior to implementation of the video service, as a direct result of the research findings.

While there are many process mapping tools that enable an understanding of work processes, FRAM is the only tool that enables variation in processes to be directly mapped. FRAM is therefore most useful for mapping processes that are non-linear, and that have many co-dependencies among tasks. A disadvantage of the method is that a FRAM can quickly become complicated and unwieldy, especially if the mapping is done at a level that is too granular for the problem at hand or if the boundaries of the system are not sufficiently constrained. Because FRAM involves mapping Work-as-Done, those who actually do the work must participate in its development; this may be a burden on resources for some organisations. Finally, while FRAM can be taught to novices such that they could produce basic models after one day of training, manipulation of the FMV software and useful interpretation of the results can be dependent on the skill and experience of the modeler.

3.2. Designing information systems to cope with unexpected events

Analysis of the thunderstorm asthma event revealed deficiencies in information systems that precluded a whole-of-system response to the emergency. In particular, information systems were found to be inadequate to support the ability to anticipate and the ability to respond. In terms of anticipation, a notification process should be developed that disseminates early information about an emerging incident to all relevant emergency management organisations. In terms of response, a centralised online system should be established to link all hospitals to ensure that they receive timely and relevant information on the medical implications of emerging events.

Using the RAG provides additional insight into the dynamic aspects of systems, particularly monitoring and anticipating, than that provided by more conventional investigation tools such as Root Cause Analysis. The RAG, however, is very dependent on the quality of the probing questions developed to assess each of the four capabilities of resilience. The questions must be designed for the specific case, and this may require specialist subject matter expertise. In addition, the combination of interviews, focus groups and ethnography required to elicit answers to the questions requires familiarity with qualitative research methods.

4. Discussion

The theory of RHC is relatively new, and tools such as FRAM and the RAG are still in their infancy. Despite this, we have learned that effective implementation in healthcare must consider that the system is dynamic, that behaviours are emergent and never wholly predictable, that causality is not knowable, and that validity of results will be limited by context. We know that local problems will impact, and will be impacted by, the larger system of which they form part, and that multiple interventions will interact, often in unpredictable ways.[9] Planning implementation and evaluation collaboratively with clinicians and patients will assist in understanding Work-as-Done and support the intervention so that it can be better matched to the needs of the workforce. Overall, RHC is showing great promise for implementation and sustainability of complex health service interventions, including digital health systems.

Teaching questions for reflection

1. What are the implications of RHC for design and sustainability of digital health systems?
2. When implementing a digital health system, how would you account for variation in everyday work?
3. What information system designs might work to improve communications during unexpected, rapid-onset, large-scale public health events?
4. How would you design large-scale information systems to incorporate the need to monitor, anticipate, respond and learn?

References

[1] R. Clay-Williams, *Restructuring and the resilient organisation: implications for health care*, in *Resilient health care*, E. Hollnagel, J. Braithwaite, and R. Wears, Editors. 2013, Ashgate Publishing Limited: Surrey, UK.
[2] C. Bieder and M. Bourrier, *Trapping safety into rules: how desirable or avoidable is proceduralization?* 2013, Surrey, UK: CRC Press.
[3] R. Clay-Williams and L. Colligan, *Back to basics: checklists in aviation and healthcare.* BMJ Qual Saf, 2015. **24**(7): p. 428-431.
[4] A.B. Haynes, T.G. Weiser, W.R. Berry, et al., *A surgical safety checklist to reduce morbidity and mortality in a global population.* N Engl J Med, 2009. **360**(5): p. 491-499.
[5] S.M. Levy, C.E. Senter, R.B. Hawkins, et al., *Implementing a surgical checklist: more than checking a box.* Surgery, 2012. **152**(3): p. 331-336.
[6] M.F. Peerally, S. Carr, J. Waring, and M. Dixon-Woods, *The problem with root cause analysis.* BMJ Qual Saf, 2017. **26**(5): p. 417-422.
[7] D. Hindle, J. Braithwaite, J. Travaglia, and R. Iedema, *Patient safety: a comparative analysis of eight inquiries in six countries*, in *Centre for Clinical Governance Research*. 2006: University of New South Wales, Sydney.
[8] L. Leape, D. Berwick, C. Clancy, et al., *Transforming healthcare: a safety imperative.* BMJ Qual Saf, 2009. **18**(6): p. 424-428.
[9] J. Braithwaite, K. Churruca, L.A. Ellis, et al., *Complexity Science in Healthcare – Aspirations, Approaches, Applications and Accomplishments: A White Paper*. 2017: Australian Institute of Health Innovation, Macquarie University, Sydney, Australia.
[10] E. Hollnagel, D.D. Woods, and N. Leveson, *Resilience engineering: concepts and precepts*. 2006, Surrey, UK: Ashgate Publishing Limited.
[11] C.S. Holling, *Resilience and stability of ecological systems.* Annu Rev Ecol Syst, 1973. **4**: p. 1-23.
[12] C. Perrow, *Normal accidents: living with high-risk technologies*. 1984, New York: Basic Books. x, 386.
[13] E. Hollnagel, J. Braithwaite, and R. Wears, *Resilient health care*. 2013, Surrey, UK: Ashgate Publishing Limited.
[14] J. Braithwaite, R.L. Wears, and E. Hollnagel, *Resilient health care: turning patient safety on its head.* Int J Qual Health Care, 2015. **27**(5): p. 418-420.
[15] P. Nilsen, *Making sense of implementation theories, models and frameworks.* Implement Sci, 2015. **10**(1): p. 53.
[16] J.D. Morecroft, *Strategic modelling and business dynamics: a feedback systems approach*. 2015, West Sussex, UK: John Wiley & Sons.
[17] J.D. Sterman, *Business dynamics: systems thinking and modeling for a complex world*. 2000, Boston, USA: McGraw-Hill.
[18] R.L. Wears, S. Perry, and A. McFauls, *Free fall—a case study of resilience, its degradation and recovery in an emergency department.* Paper presented at the Second International Symposium on Resilience Engineering, 2006: p. 8-10.
[19] M.M. Günal and M. Pidd, *Discrete event simulation for performance modelling in health care: a review of the literature.* Journal of Simulation, 2010. **4**(1): p. 42-51.
[20] K. Kotiadis, A.A. Tako, and C. Vasilakis, *A participative and facilitative conceptual modelling framework for discrete event simulation studies in healthcare.* J Oper Res Soc, 2014. **65**(2): p. 197-213.

[21] Morecroft J., *Agent-based and individual-based modelling*. 2012, New Jersey, USA: Princeton University Press.

[22] Gilbert N, *Quantitative applications in the social sciences: agent-based models*. 2008, California, USA: Sage Publications.

[23] R. Clay-Williams, M. Baysari, N. Taylor, et al., *Service provider perceptions of transitioning from audio to video capability in a telehealth system: a qualitative evaluation.* BMC Health Serv Res, 2017. **17**(1): p. 558.

[24] E. Hollnagel, *FRAM: the Functional Resonance Analysis Method: modelling complex socio-technical systems*. 2012, London, UK: CRC Press.

[25] R. Hill. *FRAM Model Visualiser: version 0.3.2*. 2016; Available from: http://functionalresonance.com/FMV/index.html.

[26] T. Pearce, *Review of response to the thunderstorm asthma event of 21-22 November 2016: final report*. 2017, Victorian Government: Melbourne, Australia.

[27] E. Hollnagel, *Epilogue: RAG - the Resilience Analysis Grid* in *Resilience engineering in practice: a guidebook*, E. Hollnagel, et al., Editors. 2011, Ashgate Publishing Limited: Surrey, UK.

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

doi:10.3233/SHTI190119

Health Behaviour Theory in Health Informatics: Support for Positive Change

Stephanie MEDLOCK[a,1] and Jeremy C. WYATT[b]
[a] *Department of Medical Informatics, Amsterdam UMC, Amsterdam Public Health Research Institute, University of Amsterdam, NL*
[b] *Wessex Institute of Health Research, University of Southampton, UK*

Abstract. The rising use of the Internet and information technology has made computerized interventions an attractive channel for providing advice and support for behaviour change. Health behaviour and behaviour change theories are a family of theories which aim to explain the mechanisms by which human behaviours change and use that knowledge to promote change. Among the best-known of these theories are the Social Learning and Social Cognitive theories, the Health Belief Model, the Theory of Reasoned Action and its successors the Theory of Planned Behaviour and the Reasoned Action Approach, and the Transtheoretical model. We discuss three examples of how behaviour change theories have been applied in computer-based interventions: a system to aid users to quit smoking, a decision aid for choice of breast cancer therapy, and an internet-based exercise program for reducing cardiovascular risk. We also discuss misapplication of theory, and reflect on how these theories can best be used. Behaviour change theory can be applied in health informatics interventions in several ways; for example, to select participants for a particular intervention, to shape the content of the intervention to effectively influence behaviour, or to tailor content to individual needs. Application of these theories to provide personalized advice ("decision support") is a young but promising area of research, and could inform other decision support interventions, including those that provide support for clinicians.

Keywords. Health behavior; Health psychology; Behaviour change theory; Theory-based design; Decision support systems, clinical

Learning objectives

After reading this chapter, the reader will be able to:

1. List some health behaviour and behaviour change theories and understand their relevance to system design and participant selection.
2. Understand the association of these models with technology adoption models and organizational change models.
3. Understand how these models relate to behaviour change techniques, and have been applied in technology-based interventions for smoking, breast cancer, and exercise to reduce cardiovascular risk.
4. Apply these models to designing an intervention for changing behaviour.

[1] Corresponding Author: S. Medlock, E-mail: s.k.medlock@amc.uva.nl.

1. Introduction to behaviour change theories

Health behaviour and behaviour change theories aim to explain the mechanisms by which (health) behaviours change, with a focus on harnessing those mechanisms to promote change [1]. These theories trace their roots to early work in the field of psychology, and B.F. Skinner's work in operant conditioning. Miller and Dollard's 1941 work on "social learning and imitation" can be considered the first behaviour change theory, asserting that people develop behavioural patterns through social interaction and reinforcement, including observing the actions and consequences experienced by others. This work formed the basis of the modern Social Learning and Social Cognitive theories.

1.1. Social Learning and Social Cognitive Theory

Social Learning Theory sought to combine the behaviourist and cognitive theories of learning, by positing that people learn through an interaction between cognitive factors, environmental influences, and behaviour. Observational learning occurs with four processes: attention (observing the modelled behaviour), retention (remembering the modelled behaviour), reproduction (attempting to imitate the behaviour), and motivation (anticipating the consequences of performing the behaviour, including social consequences). Reinforcement (external consequences) and self-control also play a role. In 1986 the theory was extended into the Social Cognitive Theory, an extensive theory of human motivation and action. In this theory, cognitive, environmental, and behavioural determinants all interact and influence one another [2]. People live and act within a social structure, which is in turn influenced by its members. Human agency can be exercised by taking action, directing others, or acting as part of a group. It encompasses intention and forethought, self-regulatory, and self-reflective mechanisms. The latter includes the important psychological construct of self-efficacy – an individual's belief that their actions can effect the desired change; that is, that they are capable of being effective in a particular task.

The Social Cognitive Theory is a general theory of behaviour, not specific to health or behaviour change. Nonetheless, it is one of the most-used theories in behaviour-change interventions, including internet-based interventions [3]. A shortcoming of the model is that it ignores the role of emotions, neurology, and physiology on behaviour. For example, behaviour often shifts as people age, without any corresponding shift in social, cognitive, or environmental influences.

1.2. The Health Belief Model

Work in the field progressed in the 1950's, as researchers sought to understand why tuberculosis screening programs had failed. The factors they elucidated formed the basis of the Health Belief Model (Figure 1).

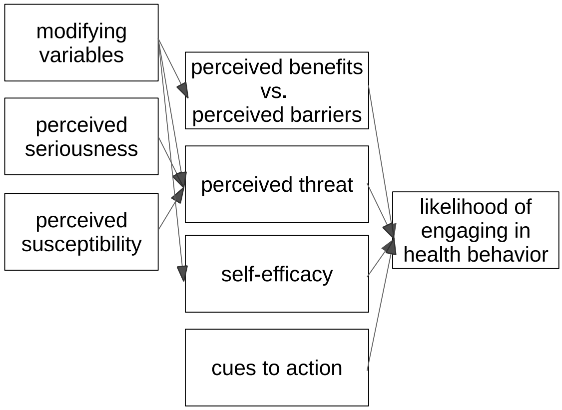


Figure 1: The Health Belief Model [4]

The authors aimed to better understand why, and under what conditions, people take action to prevent, detect and diagnose disease [4], which they termed "health behaviours" (in contrast to "illness behaviours:" the behaviours a person who perceives themselves as ill may engage in to manage or treat their condition). In the case of tuberculosis screening, most people understood that the disease was serious (perceived seriousness), but many people did not believe they were likely to catch it (perceived susceptibility). Another factor which influenced the likelihood of undergoing screening for tuberculosis was the belief that screening was effective, or the benefit of early detection – the perceived benefit of engaging in the behaviour. This was weighed against the perceived barriers, such as fear of exposure to x-rays. All of these variables were likely influenced by modifying variables such as age and social norms. The authors observed that the act of finally deciding to engage in the health behaviour is prompted by a cue – an external event that causes the behaviour to change. This could be an event that changes the perceived threat (e.g. experiencing worrying symptoms or a friend developing tuberculosis) or a public health intervention, such as a screening campaign.

Although originally intended to be a descriptive model, the Health Belief Model has also been applied both to design interventions and to predict health behaviours. The model construct "Perceived benefits and perceived barriers" has been shown to be the strongest predictor from this model [5]. A shortcoming of the Health Belief Model is its focus on individual choice, with no explicit mention of social influences or other external factors. It also assumes that health choices will be deliberate, thus ignoring unconscious choices (e.g. habit). The original model was formulated for relatively simple behaviours, such as getting a test or an inoculation. For more complex behaviours, perceived ability to perform the action (self-efficacy) is an issue. Self-efficacy was added to the model later, drawing from social cognitive theory [6].

1.3. Theory of Reasoned Action, Theory of Planned Behaviour, and Reasoned Action Approach

Shortly after publication of the Health Belief Model, Fishbein and Ajzen introduced the Theory of Reasoned Action (1967). While the Health Belief Model arose from the public health discipline, the Theory of Reasoned Action arose from social psychology and theories of attitude, particularly the Theory of Propositional Control. It asserts that *behavioural intention* (the intention to engage in a behaviour) can be predicted by the person's attitude toward that behaviour and the *subjective norm* (perceived social pressure). Attitudes about performing an act are composed of beliefs about the consequences of the act, and the subjective evaluation of (or weight given to) these consequences. The model assumes, perhaps unwisely, that behavioural intention is strongly correlated with actual behaviour [7].

The Theory of Reasoned Action is probably best known in the field of health informatics for its influence on the development of the Technology Acceptance Model (TAM)[8], which asserts that the behavioural intention of using a particular technology is predicted mainly by the user's attitude toward the technology (Reasoned Action construct "attitude"), which is in turn predicted by perceived usability and usefulness[2].

The Theory of Reasoned Action was succeeded in 1985 by the Theory of Planned Behaviour, developed by Ajzen to improve the predictive accuracy of the model by adding *perceived behavioural control* (self-efficacy), again drawn from social cognitive theory, as a construct [9]. The authors note that perceived behavioural control can influence all of the other factors in the model, including actually performing the behaviour (i.e. the individual may intend to perform the behaviour, but ultimately does not do so because they feel they cannot.)

The most recent version is the Reasoned Action Approach, an attempt to integrate the work of Fishbein, Ajzen, and several other models of behaviour change (Figure 2) [10]. It added beliefs about behaviour, norms, and control as formative constructs, and acknowledged the influence of external factors on shaping these beliefs.

Like the Health Belief Model, the Reasoned Action family of models are limited to *reasoned* action, implicitly a conscious process. It also posits that intention leads directly to action. This is often not the case; people sometimes engage in behaviours without conscious choice (e.g. habits) and frequently do not engage in a behaviour despite good intentions.

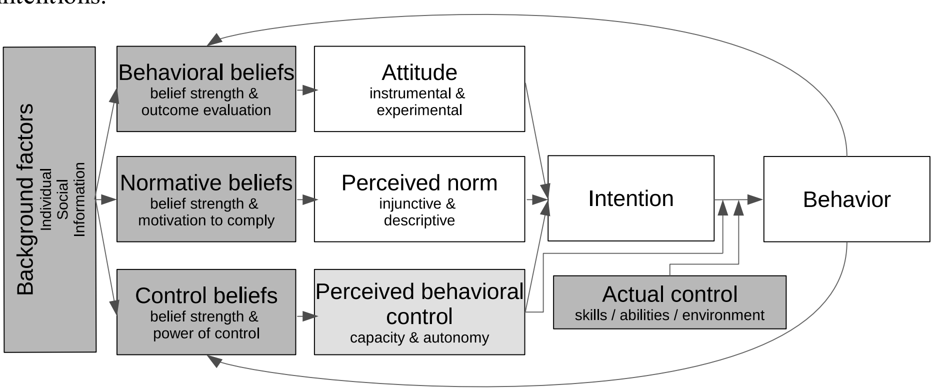


Figure 2: Reasoned Action Approach [10], showing constructs inherited from the Theory of Reasoned Action (white) and Theory of Planned Behaviour (light gray), as well as those new to the Reasoned Action Approach (dark gray).

[2] See also Chapter 6, "Technology Acceptance Models in health informatics: TAM and UTAUT".

1.4. Transtheoretical model

In 1982, shortly before the publication of Ajzen's Theory of Planned Behaviour, another highly influential model was introduced: the Transtheoretical or "stages of change" model. As the name implies, the model was developed through the integration of several behavioural and psychological models, and proposed that behavioural change occurs in five stages [11] (Box 1). A final stage, *termination*, was added later.

The authors also recognized 10 processes of change, and noted that particular cognitive processes tend to be used at different stages of change (Box 1). The verbal processes tend to play a large role in the early states of change. *Self-reevaluation* and *self-liberation* tend to come into the action phase, and *counterconditioning* and *stimulus control* bridge the action and maintenance phases. *Social liberation* plays a role in all phases. Self-efficacy and *temptation* (the strength of the desire to engage in the old behaviour) were added in later revisions of the model. The transtheoretical model has also been applied to and influenced research on organizational change – the study of preparing individuals and organizations for changes in the workplace [12]. "Resistance to change" is modelled as a mismatch between the readiness of the leadership for change and the stage of change of the employees. Studies across a range of behaviours show that before an action is taken, about 40% of people are in the pre-contemplation stage, and thus will likely resist change if the organization leadership proposes it. This can be addressed by assessing the employees' readiness to change, and taking action (individualized or collectively) according to their stage of change, for example by activities which raise awareness of the need for change for employees in the pre-contemplation phase.

Like the Health Belief Model, the Transtheoretical model portrays behaviour as mainly individual, with social influences playing only a minor role. It also assumes that people plan changes before making them. Armitage has suggested that the "five stages" can be better modelled as only two: a motivational phase (where a person prepares to change) and a volitional phase (where a person executes the change) [13] (Figure 3).

Stages of change
Precontemplation: Not yet thinking about change, may not be aware that change is needed
Contemplation: Thinking about the change
Preparation: Becoming determined to change
Action: Taking action to change
Maintenance: Maintaining the new habit
Termination: The new behaviour no longer requires active maintenance

Processes of change
Consciousness-raising: Seeking information about the behaviour
Self-liberation: Belief in the ability to change
Social liberation: Seeking and recognizing social support for the new behaviour
Self-reevaluation: Changing one's self-image in line with the new behaviour
Environmental reevaluation: Seeking and recognizing the effect of the old and new behaviour on others
Counter-conditioning: Substituting new, healthier behaviours for old habits
Stimulus control: Removing cues that trigger the old behaviour
Reinforcement management: Recognizing rewards from others and creating rewards for the new behaviour
Helping relationships: Seeking and recognizing support from others for the new behaviour

Figure 3. Armitage's two-stage model of change [13].

1.5. Translating theories to interventions

Although the above theories are among the best-known, there are many other behaviour change theories – 83 according to Davis et al [14]. Often more than one is used when designing an intervention. Michie et al. have worked extensively to aid the application of these theories by mapping elements of each theory (“constructs”) to behaviour change techniques. Their group has identified 93 distinct behaviour change techniques, such as goal-setting (setting a concrete and achievable behavioural goal), and mapped these to 14 theoretical domains [15]. These insights have been organized into the Behaviour Change Wheel and the Theoretical Domains Framework [16]. At the centre of the Behaviour Change Wheel is the COM-B model, a general model of behaviour which states that an individual will engage in a Behaviour if they have the Capability, Opportunity, and Motivation to do so. The next level of the wheel maps these to nine intervention functions (high-level techniques) and seven policy categories (reflecting policies that can facilitate the techniques). Thus, for example, "goal setting" would be classified under Motivation in the COM-B model.

A practical example of applying this framework to design an intervention is given by Mangurian et al. [17]. Their goal was to increase the rate of screening for cardiovascular disease by primary care providers in patients with severe mental illness. They followed the eight steps outlined in the Behaviour Change Wheel framework. First, focus groups were formed with providers, patients, and managers to identify barriers and potential target behaviours, to identify potential targets and select the target behaviour: ordering of metabolic screening tests (e.g. HbA1c) by primary care providers at community mental health clinics. The behaviour was then described in detail, including its place in the providers' workflow, to identify what change was needed. The researchers identified barriers in 10 behavioural constructs, e.g. "Screening for diabetes is low on the priority list for these patients" (constructs: goal setting, motivation, attitudes [Reasoned Action]) and "I don't know exactly what the guidelines recommend" (constructs: knowledge, perceived behavioural control [Reasoned Action]). These problems were then mapped to 18 intervention functions. For example, lack of knowledge could be addressed through education, or through environmental restructuring so that the knowledge is available when it is needed, e.g. through a decision support system. They also identified 8 policy strategies to support the interventions, e.g. providing training on the content of the guidelines, and 7 behaviour change techniques applicable to this situation, e.g. self-regulation (Social Cognitive Theory) in the form of feedback on their individual screening rates. Finally, they defined the mode of delivery for each intervention. The authors reported that using the framework helped them in applying the underlying behaviour change theories to their intervention.

As noted by Kok et al., designing an effective behaviour change intervention is not simple [18]. Success depends on ensuring that changing the selected determinants will result in the desired behaviour, choosing behaviour change methods that affect these determinants, and executing the methods correctly so that they can be effective. The authors propose an Intervention Mapping taxonomy and protocol. The taxonomy consists of 13 behavioural determinants with methods for changing each, the theories on which these methods are based, and evidence of their effectiveness by specifying the theoretical parameters under which the method is effective or not. The taxonomy is not limited to the individual level, but also identifies methods of change at higher ecological

Box 1: Stages and Processes of change in the transtheoretical model [10]

levels. For example, the determinant "environmental conditions" can be affected by offering technical assistance (according to the Diffusion of Innovations theory, among others). Also, change methods at the individual level can be directed toward agents at higher levels (e.g., consciousness raising), in combination with change methods for higher levels (e.g., agenda setting). The protocol for designing theory- and evidence-based behaviour change interventions consists of six steps: (1) conducting a needs assessment; (2) creating a matrix of change objectives by mapping behaviours to behavioural determinants to determine intervention targets; (3) select theory-based intervention methods that can influence the determinants, and translate these to practical interventions; (4) integrate the interventions into a programme; (5) organize adoption, implementation and maintenance of the program by identifying program users and supporters and addressing their needs; (6) create an evaluation plan to measure the effect.

2. Usage of health behaviour and behaviour change theories in health informatics

As computers and the internet have become more integrated into our lives, they have become increasingly attractive platforms for behaviour change interventions [3]. Many computer-based interventions simply provide information, with theory guiding what information is presented, to whom and in what ways. More complex interventions use specific data about the user to tailor the information that is presented or guide the user in making choices about their health, and thus can be considered a type of decision support system. Likewise, the goal of decision support is often to change behaviour – either the behaviour of health professionals on behalf of their patients, or the behaviour of patients themselves in self-management systems. Apparently well-designed decision support systems often go unused, or fail to deliver the expected effect on health or health care [19]. One possible path toward improving the success of systems is to draw from existing cognitive and behavioural theories, to determine how the system's advice should be presented to be most persuasive and most helpful to the end user.

Examples of theory-based systems described below that provide patient- or situation-specific advice to aid in making a health-related decision are the Tailored Print Smoking Cessation system, the BresDex decision aid for women with breast cancer, and the Active Living Every Day internet-based intervention for reducing cardiovascular risk.

2.1. Smoking cessation

The Tailored Print Smoking Cessation system generated person-specific advice based on the transtheoretical model. [20] The system asked the user questions based on the constructs of the transtheoretical model, including the 10 processes of change, temptations, and self-efficacy. For example, the system might ask the user to rate the statement, "I tell myself I can choose to smoke or not" (construct Self-Liberation). The system then compared the user's answers to relevant norms and used decision rules to determine which written interventional materials were appropriate for the user. The user's stage of change was determined, and then materials were selected to help move the user to the next stage. The person was reassessed every 3-6 months, and the system also incorporated data from previous assessments, and generated a feedback report. This report included a comparison of the individual's progress to a set of norms derived from data on people who ultimately were and were not successful at quitting smoking, as well as to the person's own past responses. The underlying rule base was complex, resulting in around 20,000 possible unique reports.

The system proved to be considerably more effective than providing non-tailored materials or simply allowing people to answer the questions and draw their own conclusions. Use of the system resulted in a self-reported 22-26% smoking-cessation point-prevalence (people who said they had not smoked in the last 24 hours; an intermediary outcome associated with long-term cessation) [20]. This means the system was nearly as effective at helping people quit smoking as intensive clinic-based interventions.

2.2. Choice of therapy for breast cancer

The BresDex decision aid provided a variety of different forms of information to help women with breast cancer in deciding between breast-conserving surgery with radiotherapy or mastectomy [21]. The Theory of Planned Behaviour has also been shown to predict behaviour in decisions such as choice of cancer therapy. As with other decision aids, the goal of this system was to help the patient come to a decision that is in line with her values. The authors used an extended version of the theory which included the construct "anticipated regret." The decision aid consisted of video clips of enacted patient experiences (all constructs), patient photos (attitudes and anticipated regret), videos of health professionals (attitudes, subjective norms, and perceived behavioural control), information on further treatment (attitudes and anticipated regret), a Your Surgery Options section (attitudes and anticipated regret), an It's Your Choice section (subjective norms and perceived behavioural control), a What's Next section (behavioural control), and a "decision support" tool which visually summarized the patient's responses to other items as a tally of pros and cons for each option (subjective norms and perceived behavioural control). The decision aid also offered information on the causes, types, and consequences of breast cancer, in line with the Common Sense Model of Illness Representations (a model of coping behaviour in disease). The effect of the decision aid was studied in a sample of 54 women using questionnaires administered before and after using the system. "Readiness to make a decision" improved significantly after using the system, as measured by the DelibeRATE score (increase from a mean score of 65 to 76, $p < 0.001$). However, knowledge of breast cancer did not change, as measured by a 10-point breast cancer knowledge scale (mean score = 8.3 before and 8.5 after; $p = 0.202$). Most women preferred breast-conserving surgery both before and after using the system, although slightly more preferred it after (27/54 participants before and 30/54 participants after).

2.3. Exercise

Active Living Every Day is an internet-delivered program designed to encourage exercise and reduce cardiovascular disease risk factors in sedentary adults [22]. Participants are screened for their level of readiness to change based on the Transtheoretical model. They are then matched to a (fictional) virtual participant at the same level of readiness to change. They can read about their virtual partner's progress, and receive reading material tailored to their state of change. The material includes activities such as Setting Goals (consciousness-raising), Scouting Physical Activity in Your Community (environmental re-evaluation, helping relationships), and Mall Walking (counter-conditioning, reinforcement management). Participants also keep a journal of their own activities and goals, and are periodically re-evaluated for their stage of change.

The intervention was tested in a controlled study with 32 participants, where the control group received a delayed intervention (started after the study outcomes were measured). In the 16-week program, activity increased by 1384 steps/day (p=0.03) in the intervention group, compared to 816 steps/day (p=0.14) for the control group. Waist circumference decreased in the intervention group but not in the control group (a change of -4.0 cm vs +0.6 cm; $p < 0.05$) and Coronary Risk Ratio reduced in the intervention group (from 5.1 to 4.7; p=0.04) while remaining constant in the control group (3.7; p=0.94).

3. Explanation of the success or failure of use of health behaviour and behaviour change theory in health IT systems

In a review of internet-based behaviour change interventions, Webb et al. found that use of theory was positively associated with effect size [3]. Three theories were used most often: the Theory of Planned Behaviour, Transtheoretical model, and Social-Cognitive theory. Figure 4 shows the effect sizes for studies that used these theories, as well as the effect for studies that used behaviour change theories in different ways. Using theory or predictors for participant selection was associated with a larger increase in effect size than using it for intervention design, and using it for both purposes had the greatest effect [3].

All three of the interventions described above made use of theory to develop the content of their interventions. Presenting tailored information (in other words, incorporating specific data about the user/patient to determine what kind of support should be provided) also showed a small positive effect. The Tailored Print application and the Active Living Every Day intervention are examples of this application of theory. Theory can also be applied to explain or predict observed behaviour. For example, clinicians' intention to use each of seven information sources to learn about a new drug was examined in a survey based on the Reasoned Action Approach. In this study, attitudes were shown to have a greater influence than subjective norms for this behaviour [23].

Theory can also be misapplied, e.g. by applying it out of the context in which it was developed and tested. An example of this is an attempt to use "credibility cues", based on Fogg's work on credibility, persuasion, and behaviour change, to encourage people to register as an organ donor via a website [24]. Fogg’s theory, based on many empirical studies of eCommerce websites, states that website credibility is based on the user's perception that the people behind a website are trustworthy and have relevant expertise. This means a site design should be: professional; make it easy to verify the information it contains; show that behind the site is a real organisation with people who have relevant expertise, are honest and trustworthy and can be contacted if need be; easy to use, useful and frequently updated; and that the site design avoids errors of all kinds and promotional content.

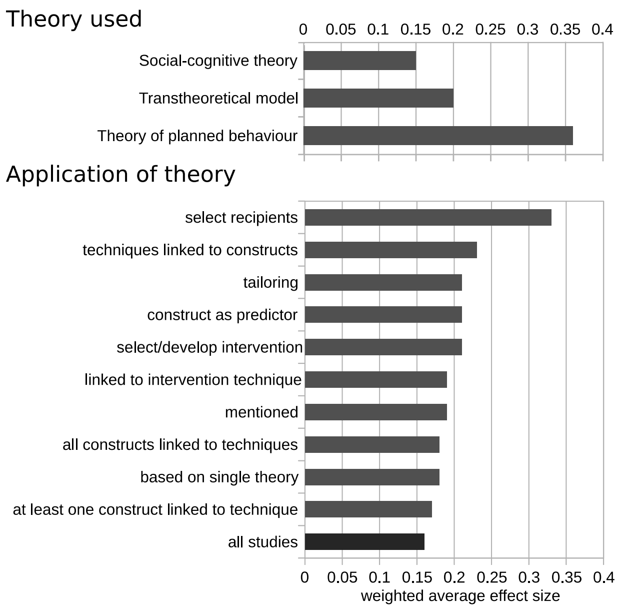


Figure 4. Effect of use of theory in internet-based behaviour change interventions. Based on data from Webb et al. [3]

To test if Fogg's theory improved sign-up rates for the NHS organ donation website, an online study was conducted in which 889 participants were randomized to use one of two versions of the website over a 3 week period, one with characteristics previously identified as signalling credibility (University logo, privacy statement, references, etc.) and one with characteristics of low credibility (advertisements, broken links, non-secure site for entering form data). This study showed no difference in the number of participants registering as an organ donor (176/466 or 37.8% in the "credible" variant, 160/423, again 37.8% in the "low credibility" variant). However, the kind of decision being made in this study – whether to allow your organs to be harvested for transplant after death – was very different from the kind of decision targeted by Fogg's persuasive technology theory – whether to purchase an item on an eCommerce website. Thus, perhaps not surprisingly, the theory was simply not valid in this context.

Another example of an intervention which did not go as planned was the Personally Controlled Health Management System for Asthma [25]. This was a web-based self-management system for asthma, with the primary goal of helping patients develop a written Asthma Action Plan. The system contained three "patient journeys," which were designed around the Health Belief Model. The first journey aimed to increase perceived susceptibility and emphasizes the importance of having a written plan in the event of an asthma attack, the second addressed perceived barriers by providing the information needed to formulate a plan, and the third addressed self-efficacy by providing encouraging tips in monthly emails. They also provided a "cue to action" by allowing patients to book a consultation online. The system was also informed by the Transtheoretical model, in that the three journeys can be viewed as appropriate for different stages of change and the social cognitive theory by incorporating social features such as polls and forums. However, only 20% of eligible participants ever logged in more than once. The primary goal of the system was to encourage developing a written Asthma Action Plan, and at the end of the study only 18% of intervention group participants had one (compared to 22% of the control group). Based on a follow-up survey, the authors cite unrecognized and unaddressed barriers as the primary reason for the poor effect of the system: for example, believing that a written plan was not necessary in their situation, being discouraged from getting one by a health care professional or by a previous negative experience with trying to create a plan, or feeling that it was unimportant compared to other priorities in their life.

Research using these theories, or theories in general, falls into two categories: studies which test a theory in an applied setting to determine if the theory applies to that setting, or studies which aim to solve a practical problem and make use of the theory to formulate a better solution [26]. The attempt to apply Fogg's theory to the NHS organ donation website can be viewed as a study of the first type. However, the outcome might be considered predictable: the target behaviour was very different than the behaviours in the studies from which the theory was derived. Once the decision has been made to become an organ donor, the user may not be deterred by being required to use an unprofessional-looking website – unlike the decision to purchase a book or music, which can always be purchased later or from another vendor. This also explains why self-efficacy needed to be added to the Health Belief Model; it was originally developed to explain simple behaviours like getting a vaccination, and needed an additional construct to be applied to more complex behaviours where the person might doubt their own ability to engage in the behaviour. Likewise, a theory may not apply if the target group is very different from the original target group or if the target behaviour is influenced by forces not accounted for in the model. For example, clinicians' behaviour is usually regulated in part by law and reimbursement policy, so it is unlikely that any behaviour change intervention would cause clinicians to behave in a way that substantially contradicted these policies. Likewise, interventions to change a health behaviour in school children or in military personnel might be quite different, since many choices in these groups are not made by the individuals themselves. A thorough review of the literature may be informative to determine if a theory has been successfully applied for that behaviour, type of person and context of interest. Several of the models discussed above have the shortcoming that they consider the behaviour of a person in isolation, disregarding social influences. These models are unlikely to apply in situations where social interaction plays a large role, or when the decision is made by a team rather than an individual. Finally, as mentioned above, correct operationalization of the theory is critical. This is likely the reason for underuse of the asthma website mentioned above: although barriers are a construct in the Health Belief Model, the investigation of barriers prior to developing the system failed to identify the barriers, which ultimately led to the system not being used. Use of a guide such as that proposed by Kok et al. can help assure that the theory is applied correctly [18].

4. Discussion

Health behaviour and behaviour change theories are widely applied in the field of psychology and have now been classified and translated into a taxonomy of practical techniques. Many health informatics interventions, particularly those directed toward healthcare professionals and patients, aim to change user behaviour in some way, and use of a behaviour change theory in intervention design or participant selection is associated with increased effectiveness in internet-based behaviour change interventions [3]. Some behaviour change theories have even been developed specifically for use in technology-based interventions [27] However, because many people working in health informatics focus on technology not psychology, the use of behaviour change theory in online or mobile interventions originates mainly from another discipline: psychologists and public health workers familiar with behaviour change theory, who are moving toward using computers and mobile devices as affordable, scalable channels to reach their patients. Many interventions developed by people working in health informatics also aim to help patients stop smoking, provide a tailored decision aid, or influence other health behaviours or health professionals, but do not report using a health behaviour or behaviour change theory. The goal of many systems designed to support clinicians could also be characterized as "behaviour change" - for example, aiming to increase screening for a particular problem, reduce ordering of unnecessary tests, or changing from the use of one medication to another [28]. One of the reasons given for failure to incorporate new evidence into practice is "habit," implying that behaviour change theory has potential application in this area as well. There is a clear opportunity for greater collaboration and multidisciplinary cooperation between the fields of health psychology, health promotion and health informatics, and an urgent need to apply health behaviour and behaviour change theories in designing and evaluating health IT interventions. One way to promote this is to include insights from psychology and behaviour change theories in health informatics education programmes.

Teaching questions for reflection

1. What do people working in health informatics need to know about behaviour change theory and techniques?
2. Does basing the design of an intervention on Behaviour Change Theory always lead to a more effective intervention? If not, why not?
3. How else can people working in health informatics make use of behaviour change theory to improve the impact of our work?
4. How can we help the field move forward and understand which Behaviour Change Theories work best in specific contexts?

References

[1] R. Davis, R. Campbell , Z. Hildon, L. Hobbs, S. Michie. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. *Health Psychology Review*. **9 (**2015), 323-344.
[2] A. Bandura. Social cognitive theory: An agentic perspective. *Annu Rev Psychol.*. **52** (2001), 1-26.
[3] T.L. Webb, J. Joseph, L. Yardley, S. Michie. Using the internet to promote health behaviour change: a systematic review and meta-analysis of the impact of theoretical basis, use of behaviour change techniques, and mode of delivery on efficacy. *J Med Internet Res*.**17** (2010) e4.
[4] I.M. Rosenstock. Why people use health services. *Milbank Mem Fund Q.* **44** (1966) S94-127
[5] I.M. Rosenstock, V.J. Strecher, M.H. Becker. Social learning theory and the health belief model. *Health Educ. Q*. **15**(1988), 175–83
[6] C.J. Carpenter. A meta-analysis of the effectiveness of health belief model variables in predicting behaviour. *Health Commun*. **25** (2010):661-9.
[7] P. Sheeran, T.L. Webb. The Intention–Behavior Gap. *Social and Personality Psychology Compass*. **9** (2016), 503-518.

[8] F.D. Davis. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology, *MIS Quarterly,* **13**(1989) 319-340.
[9] I. Ajzen. The theory of planned behavior. *Organizational Behavior and Human Decision Processes*. **50** (1991), 179-211M. Fishbein, I. Ajzen. Predicting and Changing Behavior: The Reasoned Action Approach. Psychlogy Press, New York, New York, 2010.
[10] J.O. Prochaska, C.C. DiClemente. Stages and processes of shelf-change of smoking – toward an integrative model of change. *Journal of Consulting and Clinical Psychology*. **51** (1983), 390-395
[11] J.M. Prochaska; J.O. Prochaska; D.A. Levesque. A Transtheoretical Approach to Changing Organizations. Administration and Policy in Mental Health; **28**(2001) 247.
[12] C.J. Armitage. Is there utility in the transtheoretical model? *British Journal of Health Psychology.* **14** (2009) 195–210.
[13] R. Davis, R. Campbell, Z. Hildon, L. Hobbs, S. Michie. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. *Health psychology review*, **9** (2015), 323-344.
[14] J. Cane, M. Richardson, M. Johnston, R. Ladha, S. Michie. From lists of behaviour change techniques (BCTs) to structured hierarchies: comparison of two methods of developing a hierarchy of BCTs. *Br J Health Psychol.* **20** (2015) 130–50.
[15] L. Atkins, J. Francis, R. Islam, D. O'Connor, A. Patey N. Ivers, R. Foy, E.M. Duncan, H. Colquhoun, J.M. Grimshaw, R. Lawton, S. Michie. A guide to using the theoretical domains framework of behaviour change to investigate implementation problems. *Implement Sci.* **12** (2017) 77.
[16] C.I. Mangurian, G.C. Niu, D.H. Schillinger, J.W. Newcomer, J. Dilley, M.A. Handley. Utilization of the Behaviour Change Wheel framework to develop a model to improve cardiometabolic screening for people with severe mental illness. *Implement Sci.* **12** (2017) 134.
[17] G. Kok, N.H. Gottlieb, G-J.Y. Peters, P/ Dolan-Mullen, G.S. Parcel, R.A.C. Ruiter, M.E. Fernández, C. Markham L.K. Bartholomew. A taxonomy of behaviour change methods: an Intervention Mapping approach, *Health Psychology Review*, **10** (2016) 297-312.
[18] S. Medlock, J.C. Wyatt, V.L. Patel, E.H. Shortliffe, A. Abu-Hanna. Modeling information flows in clinical decision support: key insights for enhancing system effectiveness. *J Am Med Inform Assoc.* **23** (2016) 1001-6.
[19] W.F. Velicer, J.O. Prochaska, C.A. Redding. Tailored communications for smoking cessation: past successes and future directions. *Drug Alcohol Rev*. **25** (2006), 49-57.
[20] S. Sivell, W. Marsh, A. Edwards, A.S.R. Manstead, A. Clements, G. El-wyn. Theory-based design and field-testing of an intervention to support women choosing surgery for breast cancer: BresDex, *Patient Education and Counseling* **86**, (2012), 179-188.
[21] L.J. Carr, R.T. Bartee, C. Dorozynski, J.F. Broomfield, M.L. Smith, D.T. Smith. Internet-delivered behavior change program increases physical activity and improves cardiometabolic disease risk factors in sedentary adults: results of a randomized controlled trial. *Prev Med* **46** (2008), 431-438.
[22] C.A. Gaither, R.P. Bagozzi, F.J. Ascione,Duane M. Kirking. A Reasoned Action Approach to Physicians' Utilization of Drug Information Sources. *Pharmaceutical Research* **13** (1996), 1291–1298
[23] T.J. Nind, I.W. Ricketts, F.F. Sniehotta, J.C. Wyatt, S. Cole. Do credibility cues affect the persuasiveness of a web site? Symposium: Influencing People with Information. Aberdeen. (2012)
[24] A.Y. Lau, A. Arguel, S. Dennis, S.T. Liaw, E. Coiera. "Why Didn't it Work?" Lessons From a Randomized Controlled Trial of a Web-based Personally Controlled Health Management System for Adults with Asthma. *Journal of medical Internet research*, **17** (2015) e283.
[25] R.A.C, Ruiter, K. Massar -Van Vugt, G. Kok. *Applying Social Psychology to Understanding Social Problems*. In: A. Golec de Zavala, A. Cichocka. *Social Psychology of Social Problems - The intergroup context*. Palgrave Macmillan, Basingstoke, UK, 2012.
[26] D.C. Mohr, S.M. Schueller, E. Montague, M.N. Burns, P. Rashidi. The Behavioural Intervention Technology Model: An Integrated Conceptual and Technological Framework for eHealth and mHealth Interventions. *Journal of Medical Internet Research*, **16** (2014), e146.
[27] J.C. Wyatt. Decision support systems. *Journal of the Royal Society of Medicine*, **93** (2000). 629-633.

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

doi:10.3233/SHTI190120

Control Theory to Design and Evaluate Audit and Feedback Interventions

Wouter T. GUDE [a,1] and Niels PEEK [b,c]
[a] *Department of Medical Informatics, Amsterdam UMC, University of Amsterdam, Amsterdam Public Health research institute, The Netherlands*
[b] *Division of Informatics, Imaging, and Data Science, School of Health Sciences, The University of Manchester, United Kingdom*
[c] *NIHR Greater Manchester Patient Safety Translational Research Centre, The University of Manchester, United Kingdom*

Abstract. Control theory is about the processes underlying the behaviour of self-regulating agents. It proposes that behaviour is regulated by a negative feedback loop, in which the agent compares the perception of its current state against a goal state and will strive to reduce perceived discrepancies by modifying its behaviour. Although studies in health informatics often do not report the use of this theory, the principle of a negative feedback loop underlies many applications in the field. This chapter describes how control theory fits within health informatics, discussing its role in the development and assessment of audit and feedback interventions in healthcare. Control theory has been used to synthesise evidence of audit and feedback, and to design and evaluate interventions to improve the quality of blood transfusion practice, cardiac rehabilitation, and intensive care. This has driven progress in our understanding of the underlying mechanisms of audit and feedback for improving health care, and has helped to design better interventions.

Keywords. Quality of Health Care; Continuous Quality Management; Clinical Audit; Feedback; Self-Regulation.

Learning objectives

After reading this chapter the reader will:

1. Have a broad understanding of Control Theory (CT) and its applications in behavioural sciences.
2. Understand the application of CT in the design and evaluation of health informatics (HI) interventions, such as audit and feedback; decision support; health behaviour change apps; and supervised machine learning.
3. Understand how CT has been used in scientific studies to evaluate and improve the design of audit and feedback interventions.

[1] Corresponding Author: Wouter Gude, E-mail: w.t.gude@amc.uva.nl.

1. Introduction of Control Theory

Control theory (CT) as espoused by Carver and Scheier [1] is a general approach to understanding the behaviour of self-regulating agents, which could be humans or artefacts. Its conception is usually traced to the publication of Wiener's seminal book on cybernetics – the science of feedback processes involving control or regulation of certain values within living or artificial systems [2]. Since then, CT has influenced a diverse range of fields including engineering, applied mathematics, economics, medicine, and cognitive and behavioural science. Ammons [3] used feedback processes in the context of human learning, and stated that a person's knowledge of their own performance, obtained through feedback on that performance, will affect the rate of learning and the competence level ultimately reached by that person.

1.1. Discrepancy-reducing feedback loop

The core component of CT is a negative feedback loop (Figure 1), termed negative because its function is to negate, or reduce, discrepancies between a perceived present state and a *reference value* (such as a goal state or standard). An agent perceives its current condition via an *input function*, and compares that perception against the reference value through a mechanism termed a *comparator*. If the agent observes a difference between the two values, it will attempt to reduce the discrepancy by performing a behaviour (termed the *output function*). The behaviour usually does not counter the discrepancy directly but has an impact on the *agent's environment*. This should lead to a different present condition, which in turn is perceived by the input function and compared to the reference value. This arrangement thus constitutes a closed loop of control, the overall purpose of which is to minimise deviations from the standard of comparison.

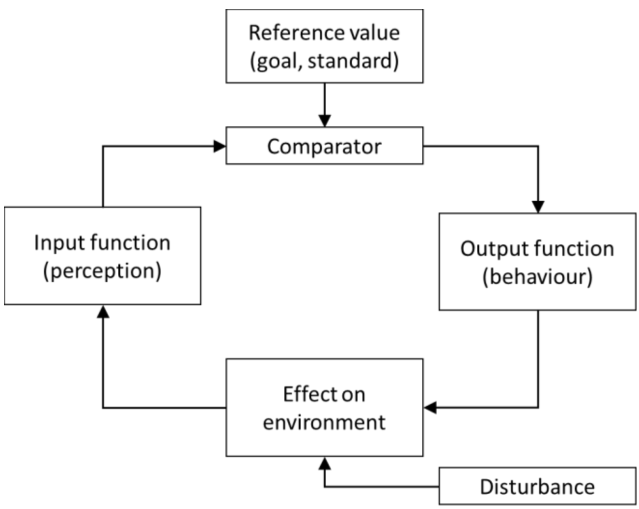


Figure 1. Negative feedback loop (from [1]).

Feedback processes like the one described above can occur in diverse physical systems; the best-known example of which is the thermostat. In this system, the input function continuously samples current air temperature from a particular environment such as a room. This input information goes to the device that compares the sensed value to the thermostat's settings. As long as there is no notable difference between values, nothing will happen. If the comparator does detect a difference between values, it sends a message that turns on the heater (output function) which begins to bring warm air into the room. The thermostat will continue to request activity from the heater until the room has warmed up enough so that it can no longer sense a discrepancy between the current air temperature and the thermostat's setting. In Section 2 of this chapter we will provide examples of feedback processes in health informatics (HI) interventions.

As Figure 1 shows, effects on a system's environment do not only depend on operations by the output function. *Disturbance*, originating outside the loop, does not affect the components of feedback loop directly, but it can modify the agent's perceptions via the input value and lead to changes in the discrepancy from the standard. These changes can be adverse (creating or increasing the discrepancy) or favourable (closing the discrepancy). In the example of the thermostat, an open window on a cold day might allow cold wind to enter the room and reduce room temperature; increasing the gap with the thermostat's target temperature range. Alternatively, a warm sun shining through the window or a large group of people standing in the room radiating excess body heat might establish the opposite effect. In that case, the result of the disturbance is that there is no need for an output adjustment because the system observes no discrepancy. Hence the main purpose of the feedback loop is not to undertake an action, but to create and maintain the perception of a specific desired condition i.e. no discrepancy between the input and reference value.

1.2. Hierarchical systems and reference values

Feedback loops are often organised in a hierarchical fashion such that there are superordinate and subordinate systems (Figure 2) [4]. Each system relates to superordinate (at the higher end of the hierarchy) or subordinate (at the lower end of the hierarchy) goals, where achievement of subordinate goals is a requisite for attaining superordinate goals. Superordinate systems act by changing the reference value of the subordinate system at the lower level in the hierarchy. That is, the output of the superordinate level sets the standards for the next lower level. In turn, the subordinate system changes the reference value for the next lower level, and so on.

Building on the thermostat example, the thermostat receives its reference value from a superordinate system which may be a person in the room. This person also has a reference value e.g. 'be comfortably warm'. Rather than operate directly on the environment to produce heat, for example, by building a fire, the person operates by providing a new reference value to the subordinate system–resetting the thermostat to a warmer temperature setting. Due to the change in reference value, the thermostat activates the furnace and the room temperature rises.

Hierarchically-organised systems each act to create their desired condition and monitor their own input which exists at their own level of abstraction. In the example of the two-level hierarchy of the person and their thermostat, the thermostat assesses air temperature whereas the person assesses their comfort level. Superordinate goals at the higher end of the hierarchy tend to be more abstract whereas subordinate goals at the lower end are more tangible and concrete. As the subordinate system executes, both systems progress towards discrepancy reduction. Higher level (i.e. abstract) goals may be achieved more gradually over time than lower level (i.e. concrete) goals.

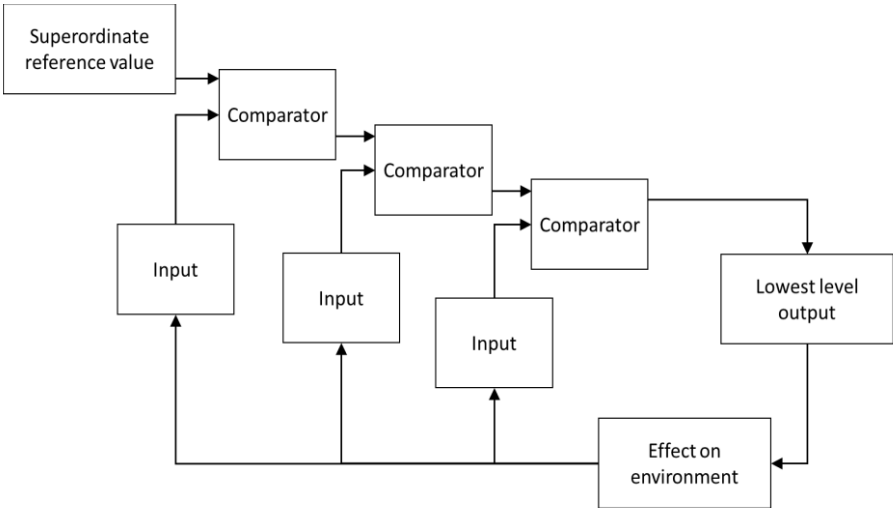


Figure 2. Three-level hierarchy of interconnected feedback loops (from [1]).

Especially amongst human behaviour theorists it is thought that hierarchies may have many levels of control [4], where behaviour outputs at the highest level are to live up to one's self-image (e.g. to be a responsible and thoughtful person) and go all the way down to muscle movements at the lowest level. So, while the basic negative feedback loop from the thermostat example is relatively simple and has limited applicability in the real world, the hierarchical approach enables modelling of arbitrarily complex systems, both in the mechanical, biological, and behavioural world.

1.3. Alternative strategies of reducing discrepancies

Feedback loops involving human behaviour are more complex than in relatively simple systems such as the thermostat. Whereas ideally potential discrepancies are resolved through behaviour, an alternative strategy is to change the reference value so that it better matches the input value. Both responses would effectively reduce perceived discrepancies, however with a different effect on the environment. For example, people being confronted by the fact that they are not achieving a lifestyle goal (e.g. walking 10,000 steps a day) sometimes respond by changing the goal rather than changing their behaviour. Other options are to reject the reference value, or to abandon the situation (physically or mentally) that signals the discrepancy, e.g. deeming the reference value unachievable or discounting the presented data or source. For instance, professionals may question the quality of underlying data when feedback on their performance indicates that it is below par. It is difficult to predict which strategy will be used in response to perceptions of discrepancies, but it is likely that alternative strategies are used when attempts by the output function to change the input seem to repeatedly fail or if the discrepancy is perceived to have a low likelihood of reduction through actions [5].

1.4. Similarities and differences with related theories

Several other theories use the concept of feedback as a central component. Key examples are goal-setting theory (Locke & Latham [6]), feedback intervention theory (Kluger & DeNisi [5]), and social cognitive theory (Bandura [7]).

Goal-setting theory [6] posits that people are motivated to achieve a goal rather than reduce discrepancies. In particular, it describes the mechanisms through which goals (comparable to CT's reference value but set by people themselves) influence behaviour and the relationship between goal characteristics and subsequent performance i.e. goal attainment. The theory proposes that specific goals are more effective than general ones (e.g. 'do your best' goals); and that challenging yet achievable goals lead to better performance than both trivial goals and overambitious goals.

Feedback intervention theory [5] considers feedback as the provision of information regarding some aspects of someone's performance on a certain task. According to this theory, people's behaviour is regulated by goals and standards which are organised, as also posited by CT, in a hierarchical fashion. Attention is limited and usually directed at an intermediate level within the hierarchy; only gaps that receive attention have the potential for change. Feedback works by providing people with new information which allows a shift of attention either toward the task or away from it. An attention shift towards the task tends to strengthen the feedback's effect on task performance whereas a shift away from it weakens the effect. The theory proposes that feedback characteristics, the nature of the task performed, and situational and personality variables determine how effectively this shift occurs.

Social cognitive theory [7] aims to guide the study of human behaviour, thought and motivation. It proposes that environment, behaviour, and personal and cognitive factors all interact as determinants of each other. The theory argues that self-efficacy, the beliefs regarding one's capabilities of successfully completing tasks, determine what challenges people choose to undertake, how much effort to expend in the endeavour, how long to persevere in the face of obstacles and failures, and whether failures are motivating or demoralising [7]. The relationship between those beliefs and behaviour is described, similar to CT, as a reciprocal learning process in which people select, react to, and learn from experiences.

2. Usage of Control Theory in health informatics

There are numerous examples of HI interventions that are based on the principle of a negative feedback loop, although few would explicitly reference CT. So, most references to CT in HI are implicit, and developers of interventions would often not consciously use the negative control loop when they design their tool or software – but the control loop would tacitly play a role in their intervention. The same holds for evaluation studies of interventions that build on CT: such evaluation studies would often assume a feedback loop around which the study is designed, without explicitly referencing CT. It is therefore challenging to assess how broadly CT is used in HI. However, we believe that there is a profound influence of CT on HI, and we will illustrate that by describing various broad areas that involve a feedback loop as a key component. The principal area that we will use to describe the usage of CT in HI is audit and feedback, but we also highlight several other areas.

2.1. Audit and feedback

Audit and feedback (A&F) interventions [8] aim to improve the quality of care by comparing observed quality parameters (quality indicators) with predefined quality targets or benchmark values. Typically, feedback on quality indicators is delivered to healthcare professionals on a regular basis, thus enabling multiple cycles through the control loop. For instance, in the Netherlands all 32 teaching intensive care units (ICUs) and 51 non-teaching ICUs participate in the National Intensive Care Evaluation (NICE) [9], and receive biannual feedback reports on standardised mortality ratios, readmission rates, length of stay, and other quality indicators. A&F is one of the most widely-used interventions in quality improvement and implementation research. It is generally used when the patient is not present, thereby making it distinctly different from clinical decision support.

We can map the components of the feedback loop (Figure 1) to elements of A&F interventions as follows. The input function (perception) consists of the feedback on clinical performance that sits at the heart of each A&F intervention, and would typically materialise through recurring, paper-based or electronic reports issued by a national audit or governing body. Feedback reports summarise the performance of individual clinicians or clinical units over a set period of time (e.g. the last 3 months) using pre-defined indicators of clinical quality, typically using a combination of graphical and numerical information (scores). Reference values may be either explicitly provided or left implicit, and different for each quality indicator. Feedback on clinical processes is often determined by reference values provided by national guidelines, while feedback on outcomes would often be determined through benchmarking between care providers. For instance, in the NICE feedback report, outcome statistics are compared to the national average and the average of a group similar sized ICUs. In A&F interventions it is commonly left to the recipient of the report (i.e. the clinician or clinical unit) to interpret the information and translate it into behaviour. For instance, a unit might decide to start a quality improvement initiative based on poor performance scores in feedback reports. If that initiative is effective, improved performance should transpire in subsequent reports and can inform the decision about whether or not to continue the programme. However, there may also be disturbances (e.g. organisational barriers) that impede actual improvements to care quality, despite the efforts of the quality improvement initiative.

Over the last decade, A&F interventions have increasingly moved from using static, paper-based (or PDF) documents to interactive electronic tools. The interactive computer interface of an e-A&F intervention may allow users to filter, drill down and further explore their performance summaries. For example, NICE participants can also view these data, updated after each monthly data upload, on a website called *NICE Online* and perform subgroup analyses [10]. In general, if an A&F system is linked to an electronic health record database, performance summaries may be generated on demand at each point in time, thus creating more flexibility for users. A recent review of e-A&F evaluation studies [11] classified them using the theoretical domains framework [12], an integrated theoretical framework synthesised from 128 theoretical constructs from 33 theories judged most relevant to implementation questions. The review found that the domains of knowledge; motivation and goals; and 'social influences' were most commonly targeted by these interventions. In contrast, professional identity and emotion were never targeted.

Despite the clear roots of A&F in CT, it is uncommon that CT is explicitly mentioned in studies concerning the design or evaluation of A&F interventions – but it does happen that related theories are mentioned. Colquhoun and colleagues [13] conducted a systematic review of the use of theory in randomised controlled trials of A&F interventions. They found that only 20 out of 140 studies (14%) reported use of theory in any aspect of the study design, measurement, implementation or interpretation. Only 13 studies (9%) reported the use of a theory to inform development of the intervention. A total of 18 different theories across educational, psychological, organisational and diffusion of innovation perspectives were mentioned. Arguably, many of these resonate with elements of CT. For instance, Social Cognitive Theory also proposes that feedback processes drive behaviour; and the Theory of Planned Behaviour postulates that attitudes, social norms, self-efficacy, and controllability provide reference values for behaviour[2].

2.2. Other uses of Control Theory in health informatics

There are other HI areas that draw upon the negative feedback loop depicted in Figure 1 and thus have roots in CT. As with the A&F literature, explicit references to CT are rare – but we would nevertheless argue that there is a clear relationship. We provide three examples here.

First, a large number of decision support systems for the management of long-term conditions such as diabetes and cardiovascular diseases have been developed that deploy a negative feedback loop for controlling important clinical parameters. For instance, Athena [14] is a clinical decision support for the management of hypertension that issues an alert to its clinician users whenever a patient's latest blood pressure measurement is too high. Similarly, Pandit [15] is a web-based diabetes management system for patients that asks them to measure and enter their blood glucose level. Whenever a glucose value is outside the normoglycaemic range, the system responds by suggesting adjustments to the patient's insulin dose. Similar mechanisms have been used in expert systems for critical care [16].

Second, many smartphone apps that assist in health behaviour change provide users with feedback on their achievements against pre-set goals, such as the number of steps taken, or time spent on physical activity per day. The feedback aims to incentivise users to increase their level of healthy behaviour when it is below target, and maintain it when it is on par.

Third, CT plays a central role in supervised machine learning methods such as Hebb's learning rule [17]; the Newton-Raphson algorithm [18]; gradient boosting [19]; and deep learning [20]. Essentially, each of these methods utilise the negative feedback loop to derive a model of an input-output function from training data. Initially, a default or random model is chosen that bears no relationship to the data, and that model is subsequently 'trained' to better fit the data. The feedback is always derived from discrepancies between observed outputs (in the data; typically called *training labels*) and predicted outputs (predicted by the model). At each iteration of the feedback loop the classifier will better approximate the input-output function that produced the data, and the discrepancies will disappear after which the process is terminated.

3. Explanation of success or failure of audit and feedback

Despite being commonly applied as a healthcare quality improvement strategy, A&F interventions yield variable and often only marginal effects [8]. Moreover, over four decades of research in the field seems to have failed to enable A&F researchers to successfully enhance intervention designs and achieve larger effects consistently. It has

---

[2] See also Chapter 4, "Assessing technology success and failure using Information Value Chain Theory", where Information Value Chain theory has been applied to A&F interventions.

been only recently that the use of extant theory has been recognised as an essential component of both the design and evaluation of A&F interventions [21]. In response there have been various studies making explicit use of CT to enhance understanding the A&F's underlying mechanisms in improving healthcare and quality interventions. We have selected four recent studies that jointly illustrated the breadth of activities that can be supported with CT. These activities include interpreting published literature; designing new interventions; secondary analysis of clinical trial data; and the design of new scientific experiments.

3.1. Synthesising evidence from A&F interventions

In an illustration of how theory can be used to synthesise published evidence from behaviour change interventions, Gardner et al. [22] used CT to organise, understand and synthesise evidence relating to behaviour change techniques within A&F. Using CT as conceptual framework, the authors hypothesise that A&F may be enhanced through the use of specific *performance targets* to permit comparison between current and target performance, and *action plans* to inform behavioural adjustment to reduce discrepancy [22]. The authors conducted a re-analysis of the 2006 Cochrane review [23], recoding each study included in the review, to test target-setting and action plans as effect-modifiers of A&F. The results however were inconclusive because very few studies explicitly described their use of targets or action plans. When Ivers et al. updated the Cochrane review in 2012 [8], and repeated Gardner's analysis, explicit targets and action plans were found to be significant effect modifiers of A&F.

3.2. Improving the design of A&F to increase uptake of evidence-based blood transfusion practice

The second illustration is a study by Gould et al. [24] that used CT to enhance the content of a feedback intervention for improving blood transfusion practice. The authors describe the feedback loop as a dynamic, iterative process of control in which "individuals manage their behaviour by knowing what they want to do or achieve (i.e. setting a goal or standard), trying to do it (i.e. action), monitoring the behaviour (i.e. audit), assessing whether they are making progress towards the goal (i.e. feedback, which informs the nature and extent of any discrepancy between behaviour and goals), and adapting what they do in light of the feedback (i.e. action planning)" [24, page 2]. They also used the taxonomy of behavioural change techniques [25] for identifying and describing intervention components that are consistent with CT and that may enhance practice. A number of the techniques included in the taxonomy encompass strategies proposed in CT, such as 'goal setting'; 'feedback on behaviour'; 'discrepancy between behaviour and goal'; and 'action planning'. The authors then aimed to enhance content in feedback documents by incorporating behavioural change techniques consistent with CT that were previously absent. For example, to incorporate goal setting as a change technique, the authors added an introductory statement in documents that proposes an evidence-based goal, e.g. "XX% of patients with [XX clinical attributes] are likely to require transfusion and so we suggest that, within your clinical team, you make this your explicit goal". The authors propose that such enhanced feedback has the potential to facilitate the enactment of CT's feedback processes and lead to larger improvements.

3.3. Quantitative process evaluation of an A&F intervention in cardiac rehabilitation

The third study illustrates how CT can be used to enrich quantitative process evaluations of A&F interventions. In this case CT was instrumental to understand the outcomes observed in a cluster randomised trial of e-A&F to improve cardiac rehabilitation [26]. The intervention involved local quality improvement teams receiving quarterly web-based feedback in combination with outreach visits. Feedback was given on 18 quality indicators and included benchmark comparisons. During the visits teams reviewed their feedback and selected indicators they wished to improve upon into their action plan–within the same web-based system. For each indicator that teams targeted for improvement, they were prompted to describe the problem, goal, and concrete actions on how to achieve the goal. During each outreach visit (corresponding to feedback cycles), teams reviewed the new performance scores and updated their action plan accordingly. The associated trial did not show any significant changes in either care processes or patient outcomes [27]. Following CT, the intervention's ineffectiveness must have been either the result of the possibility that feedback indicating sub-benchmark performance still failed to convince recipients to change, actions were not completed, or completed actions were ineffective. We designed a two-part study to investigate the first gap. Part 1 was a laboratory experiment involving 41 individual cardiac rehabilitation professionals who were given two feedback reports in an adjusted version of the web-based feedback system. These professionals were asked to select indicators for improvement, based on the feedback. If their response was at odds with CT's hypothesis (i.e. indicator's performance score was below the benchmark but not selected for improvement; or the score was above benchmark and still selected for improvement), they were asked to explain their choice. Part 2 was a field study concerning a secondary analysis of the trial data, in which multidisciplinary teams selected indicators for improvement across multiple cycles of feedback. Regression techniques were applied to assess determinants of cardiac professionals' intentions to improve practice. The principal findings were that performance scores and benchmark comparisons influenced intentions, but between one third and half of the time intentions were at odds with CT because professionals either disagreed with benchmarks; deemed improvement unfeasible; or did not consider the indicator an essential aspect of care quality. In addition, it revealed that intentions remained similar in subsequent feedback cycles (because actions were not completed) and that professionals prioritised improving data quality rather than care quality. This study contributed to the understanding of A&F in both the current intervention and in general, in the sense that it quantified how often the feedback loop stagnates and provided insight into the determinants and reasons for not following feedback.

3.4. Understanding the influence of A&F in pain management in intensive care

In the previous example CT was used as a conceptual framework in a post-hoc analysis of decision processes. The final illustration also concerns work of our own research team and builds on the findings of the previous study, but in this case CT was used in the very design of the experiment. This study involved an e-A&F dashboard providing intensive care teams with periodic feedback on four pain management indicators [28,29]. Inspired by the cardiac rehabilitation study we recognised that healthcare professionals often already have beliefs about their clinical performance and feedback may fail to change those beliefs. We studied the extent to which those beliefs correspond to actual practice; how they are influenced by feedback; and, ultimately, how feedback changes intentions to improve practice. To that end we designed an online two-step experiment, driven by CT, to elicit these beliefs and intentions before and after receiving first-time feedback. The experiment took place upon first login into the dashboard; 83 intensive care professionals from 21 units participated. In step 1, professionals were presented with the indicator descriptions whilst withholding all performance feedback (that is, no performance scores or benchmark comparisons were displayed). Professionals were asked to estimate for each indicator their own unit's performance score, the national average score, the minimum score they would consider "good performance" (target), and whether or not they would perform actions to improve. The study found that half of the time professionals overestimated their own performance and rarely underestimated it. Targets were set very high. In step 2 professionals received feedback on their performance. Feedback included the unit's own performance, median and top 10% peer performance, and improvement recommendations based on peer comparisons (good performance; room for improvement; or improvement recommended). Professionals were asked again, but now given the performance information at hand, what their performance target was and whether they intended to improve practice. If improvement intentions were at odds with CT in step 1 (score < target and no intention; or score ≥ target and still intention) or in step 2 (e.g. room for improvement but no intention to improve) we asked professionals to explain their choice. Also, if there were discrepancies between intentions in the first and second step, professionals were asked what feedback elements drove them to change (e.g. measured score or benchmark was higher/lower than expected). Even before receiving any feedback some 68% of professionals' intentions corresponded with the feedback recommendations. In other words, while professionals were not very good at estimating absolute performance, they had good intuitions about whether it was on target or not– without seeing any numerical information. After receiving the feedback, this number increased to 79%. In more than half of the cases in which units were already top performers, professionals still wanted to improve. In 8% of cases professionals lacked improvement intentions because they did not consider the indicators important; did not trust the data; or deemed benchmarks unrealistic. This research concluded that audit and feedback does indeed help healthcare professionals to work on those aspects for which improvement is recommended because it increases the accuracy of their clinical performance perceptions. However, given the abundance of professionals' prior good improvement intentions, efforts to optimise A&F interventions should focus on translating those intentions into (effective) actual change in clinical practice.

4. Discussion

Control theory (CT) provides a conceptual framework for self-regulation and human behaviour and has already demonstrated its usefulness for the field of HI and in particular A&F interventions. In the A&F literature CT has been used to synthesise evidence of interventions, enhance their design, explain why interventions were or were not successful, and generate hypotheses about how feedback mechanisms work in practice. Nevertheless, the majority of studies have not explicitly reported the use of CT (or other relevant theories) for such purposes.

The simplicity of CT's negative feedback loop makes for an elegant framework that is widely applicable, but it also has limitations. Individuals may compare feedback to multiple internal standards or goals at the same time; based on beliefs about past performance, expectations, norms, or an ideal goal [5]. Further, HI interventions like A&F interventions are typically complex and placed into a social and organisational context. This context is not in the scope of CT; taking it into account would require the use of different theories such as social cognitive theory. Finally, in contrast to for example feedback intervention theory or goal setting theory, CT provides no guidance as to which factors related to the context, recipients, or feedback itself may influence success of the feedback loop. Nevertheless, it is fair to say that CT has been very influential in our thinking about information systems and behaviour and will undoubtedly continue to do so.

Teaching questions for reflection

1. Explain the various components of the negative feedback loop proposed by CT and their role in A&F.
2. What are, according to CT, the four possible responses a physician might give after being confronted by feedback that indicates that their clinical performance is below average?
3. Design a two-level hierarchical system reflecting a physician at the lower level using a decision support system at the higher level and explain how they interact.
4. Describe three CT-derived hypotheses one could test in a trial setting to increase A&F effectiveness.

References

[1] C.S. Carver, and M.F. Scheier, Control theory: a useful conceptual framework for personality-social, clinical, and health psychology., *Psychol. Bull.* **92** (1982) 111–135. doi:10.1037/0033-2909.92.1.111.
[2] N. Wiener, Cybernetics: Control and communication in the animal and the machine. Cambridge, MA: MIT Press, 1948.
[3] R.B. Ammons, Effects of Knowledge of Performance: A Survey and Tentative Theoretical Formulation, *J. Gen. Psychol.* **54** (1956) 279–299. doi:10.1080/00221309.1956.9920284.
[4] W.T. Powers, Feedback: Beyond Behaviorism, *Science (80-.).* **179** (1973) 351–356. doi:10.1126/science.179.4071.351.
[5] A.N. Kluger, and A. DeNisi, The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory., *Psychol. Bull.* **119** (1996) 254–284. doi:10.1037/0033-2909.119.2.254.
[6] E.A. Locke, and G.P. Latham, Building a practically useful theory of goal setting and task motivation. A 35-year odyssey., *Am. Psychol.* **57** (2002) 705–717. doi:10.1037/0003-066X.57.9.705.
[7] A. Bandura, Social foundations of thought and action : a social cognitive theory, 1986.
[8] N. Ivers, G. Jamtvedt, S. Flottorp, J.M. Young, J. Odgaard-Jensen, S.D. French, M.A. O'Brien, M. Johansen, J. Grimshaw, and A.D. Oxman, Audit and feedback: effects on professional practice and healthcare outcomes., *Cochrane Database Syst. Rev.* **6** (2012) CD000259. doi:10.1002/14651858.CD000259.pub3.
[9] N. van de Klundert, R. Holman, D.A. Dongelmans, and N.F. de Keizer, Data Resource Profile: the Dutch National Intensive Care Evaluation (NICE) Registry of Admissions to Adult Intensive Care Units., *Int. J. Epidemiol.* **44** (2015) 1850–1850h. doi:10.1093/ije/dyv291.
[10] N.F. de Keizer, L. Peute, E. van der Zwan, M. Jaspers, and E. de Jonge, NICE Online; a web-based tool for monitoring performance measures in intensive care, *Netherlands J. Crit. Care*. **15** (2011) 131–136.
[11] T. Tuti, J. Nzinga, M. Njoroge, B. Brown, N. Peek, M. English, C. Paton, and S.N. van der Veer, A systematic review of use of behaviour change theory in electronic audit and feedback interventions, *Implement. Sci.* (2017).
[12] C. James, O. Denise, and M. Susan, Validation of the theoretical domains framework for use in behaviour

change and implementation research, *Implement. Sci.* **7** (2012) 37. doi:10.1186/1748-5908-7-37.

[13] H.L. Colquhoun, J.C. Brehaut, A. Sales, N. Ivers, J. Grimshaw, S. Michie, K. Carroll, M. Chalifoux, and K.W. Eva, A systematic review of the use of theory in randomized controlled trials of audit and feedback., *Implement. Sci.* **8** (2013) 66. doi:10.1186/1748-5908-8-66.

[14] M.K. Goldstein, B.B. Hoffman, R.W. Coleman, M.A. Musen, S.W. Tu, A. Advani, R. Shankar, and M. O'Connor, Implementing clinical practice guidelines while taking account of changing evidence: ATHENA DSS, an easily modifiable decision-support system for managing hypertension in primary care., *Proceedings. AMIA Symp.* (2000) 300–4. http://www.ncbi.nlm.nih.gov/pubmed/11079893.

[15] A.C.R. Simon, F. Holleman, W.T. Gude, J.B.L. Hoekstra, L.W. Peute, M.W.M. Jaspers, and N. Peek, Safety and usability evaluation of a web-based insulin self-titration system for patients with type 2 diabetes mellitus, *Artif. Intell. Med.* **59** (2013). doi:10.1016/j.artmed.2013.04.009.

[16] L. Fagan, and J. Kunz, Extensions to the Rule-Based Formalism for a Monitoring Task, in: Buchanan BG, Shortliffe EH, "Rule-Based Expert Syst. MYCIN Exp. Stanford Heuristic Program. Proj., Addison-Wesley, 1984: pp. 397–423.

[17] F. Attneave, M. B., and D.O. Hebb, The Organization of Behavior; A Neuropsychological Theory, *Am. J. Psychol.* **63** (1950) 633. doi:10.2307/1418888.

[18] T.J. Ypma, Historical Development of the Newton–Raphson Method, *SIAM Rev.* **37** (1995) 531–551. doi:10.1137/1037125.

[19] J.H. Friedman, Greedy function approximation: A gradient boosting machine, *Ann. Stat.* **29** (2001) 1189–1232. doi:DOI 10.1214/aos/1013203451.

[20] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, 2016.

[21] N.M. Ivers, A. Sales, H. Colquhoun, S. Michie, R. Foy, J.J. Francis, and J.M. Grimshaw, No more "business as usual" with audit and feedback interventions: towards an agenda for a reinvigorated intervention., *Implement. Sci.* **9** (2014) 14. doi:10.1186/1748-5908-9-14.

[22] B. Gardner, C. Whittington, J. McAteer, M.P. Eccles, and S. Michie, Using theory to synthesise evidence from behaviour change interventions: The example of audit and feedback, *Soc. Sci. Med.* **70** (2010) 1618–1625. doi:10.1016/j.socscimed.2010.01.039.

[23] G. Jamtvedt, J.M. Young, D.T. Kristoffersen, M.A. O'Brien, and A.D. Oxman, Audit and feedback: effects on professional practice and health care outcomes., *Cochrane Database Syst. Rev.* (2006) CD000259. doi:10.1002/14651858.CD000259.pub2.

[24] N.J. Gould, F. Lorencatto, S.J. Stanworth, S. Michie, M.E. Prior, L. Glidewell, J.M. Grimshaw, and J.J. Francis, Application of theory to enhance audit and feedback interventions to increase the uptake of evidence-based transfusion practice: an intervention development protocol., *Implement. Sci.* **9** (2014) 92. doi:10.1186/s13012-014-0092-1.

[25] S. Michie, M. Richardson, M. Johnston, C. Abraham, J. Francis, W. Hardeman, M.P. Eccles, J. Cane, and C.E. Wood, The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions., *Ann. Behav. Med.* **46** (2013) 81–95. doi:10.1007/s12160-013-9486-6.

[26] W.T. Gude, M.M. van Engen-Verheul, S.N. van der Veer, N.F. de Keizer, and N. Peek, How does audit and feedback influence intentions of health professionals to improve practice? A laboratory experiment and field study in cardiac rehabilitation., *BMJ Qual. Saf.* **26** (2017) 279–287. doi:10.1136/bmjqs-2015-004795.

[27] W.T. Gude, M.M. van Engen-Verheul, S.N. van der Veer, H.M.C. Kemps, M.W.M. Jaspers, N.F. de Keizer, and N. Peek, Effect of a web-based audit and feedback intervention with outreach visits on the clinical performance of multidisciplinary teams: a cluster-randomized trial in cardiac rehabilitation., *Implement. Sci.* **11** (2016) 160. doi:10.1186/s13012-016-0516-1.

[28] W.T. Gude, M.J. Roos-Blom, S.N. van der Veer, E. de Jonge, N. Peek, D.A. Dongelmans, and N.F. de Keizer, Electronic audit and feedback intervention with action implementation toolbox to improve pain management in intensive care: Protocol for a laboratory experiment and cluster randomised trial, *Implement. Sci.* **12** (2017) 68. doi:10.1186/s13012-017-0594-8.

[29] W.T. Gude, M.-J. Roos-Blom, S.N. van der Veer, D.A. Dongelmans, E. de Jonge, J.J. Francis, N. Peek, and N.F. de Keizer, Health professionals' perceptions about their clinical performance and the influence of audit and feedback on their intentions to improve practice: a theory-based study in Dutch intensive care units, *Implement. Sci.* **13** (2018) 33. doi:10.1186/s13012-018-0727-8.

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

doi:10.3233/SHTI190121

Implementing and Embedding Health Informatics Systems – Understanding Organizational Behaviour Change Using Normalization Process Theory (NPT)

Mike BRACHER[a,1] and Carl R. MAY[b]
[a] *School of Health Sciences, University of Southampton, Southampton, UK*
[b] *Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, UK*

Abstract. Successful implementation of health informatics systems depends not only on efficient performance of intended tasks, but also integration into existing working relationships and environments. Implementation is an understudied area in health informatics research, and relevant empirical evidence is often absent from strategic decision making. Implementation theories such as Normalization Process Theory (NPT) can help address this gap by providing explanations for relevant phenomena, proposing important research questions, and framing collection and analysis of data. NPT identifies, characterizes, and explains mechanisms that have been empirically demonstrated to affect implementation processes and outcomes. These explanations are generalizable and facilitate comparative investigations. The first section of this chapter introduces the four main constructs of NPT (coherence, cognitive participation, collective action, and reflexive monitoring) and their constituent components. Each component is discussed with reference to a real-world example, and relationships between the four constructs are explored. The second section explores how NPT has been applied in both prospective planning of interventions and their evaluation, as well as retrospective exploration of factors promoting or inhibiting successful implementation. We examine two examples from published literature: firstly, prospective planning of an evaluation study on implementation of a digital health intervention for Type-2 diabetes; and secondly an evaluation of implementation of a new electronic preoperative information system within a surgical pre-assessment clinic. The chapter concludes with reflections on some limitations of NPT as a theoretical framework.

Keywords. Implementation science, Process evaluation, Organizational behavior change, Change management, Developer-user co-design

Learning objectives

After reading this chapter, the reader will be able to:

1. Understand the basic NPT framework, and describe the four main constructs;
2. Be familiar with example applications of NPT relevant to health informatics;
3. Understand how to apply NPT in prospective planning and evaluation of implementation of health informatics systems.

[1] Corresponding Author, Mike Bracher, E-mail: M.J.Bracher@soton.ac.uk

1. Introduction to Normalization Process Theory (NPT)

'Implementation theories are useful. They provide explanations for relevant phenomena, propose important research questions, and frame the collection and analysis of data. These explanations are generalizable, and facilitate comparative studies. Implementation researchers now have a wide range of useful theoretical tools at their disposal…Normalization Process Theory (NPT), is one of these. It identifies, characterizes and explains mechanisms that have been empirically demonstrated to motivate and shape implementation processes and affect their outcomes.' [1]

Why are new technologies and working practices implemented successfully in some settings, but not in others? What affects whether a new technology or practice will be implemented in the first place, and whether it will 'stick' in the longer term (that is, become incorporated into routine work within an organization)? NPT has developed from empirical attempts to answer such questions[2]. In this first section, we will explore the theory in terms of its main constructs and their components, to understand how NPT provides a framework for understanding implementation[3].

1.1 Understanding implementation as a set of processes

NPT focuses on *action* (both individual and social) - that is, what people *do*, rather than what they say or think – and on the *processes* through which these actions take shape. NPT is grounded in the premise that implementation of an e-health or informatics application involves human actors in four things: (i) changes in goal-directed interactions with material and virtual things (physical infrastructure, hardware and software); (ii) relational restructuring (changes in the experience and organization of human relations); (iii) normative restructuring (changes in the rules and resources that make action possible); and (iv) organizing logics (changes in the ways that whole systems are defined and understood). From these stem *specific kinds of work* and it is from these that implementation processes are derived [2]. These are: *Coherence, Cognitive Participation, Collective Action,* and *Reflexive Monitoring*. Each construct has four sub-components, which set out more specific kinds of work that occur within each domain. The remainder of this section will describe these elements and their relationship within the overall framework (readers may also view the Appendix which illustrates the relationship between constructs and their components as tables).

1.2 Making sense of new technologies and practices (Coherence).

Coherence relates to 'the *sense-making work* that people do individually and collectively when they are faced with the problem of operationalizing some set of practices' [3]. Sense making is the work that people do to understand that the intervention and its associated practices.

---

[2] Readers wishing to further explore the history and context of NPT development may refer to the 'Background' section of May et al.'s (2018) systematic review of NPT use in feasibility studies and process evaluations [1],

[3] This chapter will not discuss relations between NPT and other theories of implementation or organizational and/or behavioural change, as it is beyond the scope of the article. Readers interested in further comparative exploration of NPT in relation to other such theories may refer Moullin et al.'s (2015) systematic review of implementation frameworks relating to innovations in healthcare [16].

1.2.1 *How is what is being implemented different from what already happens? (Differentiation)*

Differentiation refers to the work that people do to understand how a new ensemble of practices is different from what came before. For example: a group of clinicians implementing a video conferencing system to interact with patients will likely want to understand how this new practice and its objects (i.e. video consultation equipment) operate differently from in-person consultation (both in terms of clinician patient interaction, and how new objects interact with other existing tools, protocols, and diagnostic implements) [3].

1.2.2 *What does the intervention mean for team working? (Communal Specification)*

Healthcare is commonly delivered by teams, many of which involve clinicians and other professionals with different skill sets and contributions to overall patient pathways. This activity involves team working, including both immediate collaboration within the same physical setting, and remote working between individuals and teams based at different locations. Introducing new technologies and ways of working therefore has the potential to change working relationships. Making sense of interventions therefore also involves questions such as, 'what needs to be done, by who, and when?' [3]. For example, the rationale for changes to information recording systems may be to reduce workloads and/or improve accuracy and responsiveness. However, this also implies changes to the tasks and divisions of labour associated with record keeping. If we imagine the hypothetical case of a new electronic health record (EHR) being implemented into a General Practice (GP) clinic, questions around *Communal Specification* might include: how will the new system change who records and/or retrieves patient information? Are these changes the same for all information pathways within the service (e.g. does it affect the pathway of care for patients managing asthma in the same way as for patients managing a mental health condition)?

1.2.3 *What does the intervention mean for specific people? (Individual Specification)*

Making sense of interventions also has an individual component; that is, how interventions will affect the tasks and responsibilities of specific people. For example, staff recruiting patients to a clinical trial need a strong understanding of the work required to secure informed consent from patients (i.e. how the conditions of a new trial will affect their specific tasks and responsibilities in recruitment) [3].

1.2.4 *How do participants see the value of the intervention? (Internalization)*

While understanding the practical aspects of the intervention (i.e. what is new, what it means for team working as well as individual responsibilities) is vital, it is also important that participants see the *value* of what is being implemented. Returning to the example of a video conferencing system for remote consultations, we might ask *how* clinicians involved in its implementation come to see its worth (or not) [3].

1.3 Establishing relationships and divisions of labour to support the intervention (Cognitive participation).

Cognitive participation refers to ‘the relational work that people do to build and sustain a community of practice around a new technology or complex intervention’[2]. While making sense of an intervention is a necessary step, successful implementation also requires that participants use this knowledge to establish responsibilities and divisions of labour that will support it. The components of *cognitive participation* point to the more specific sub-types of work that take place within this domain

1.3.1 *Who are the key people and what are they doing? (Initiation)*

Implementation of new technologies or practices in healthcare services is often delegated to a small group of managers and professionals [3]. These people frequently take the lead in setting up systems, procedures, and protocols, as well as engaging with others involved in implementation to ensure that necessary actions are undertaken. This construct draws our attention to questions of process: how have key people been identified? How has their role been established?

In the case of the EHR implementation within a GP surgery, we expect that (at least) four kinds of key people will exist: clinicians (who retrieve information for the purpose of providing treatment and care); administrators (who provide support to clinicians through information work); specialist health informatics and IT staff (who support implementation with specialist knowledge and skills); and patients (who are directly and indirectly interacting with this system as those move along pathways of care and treatment). Our focus here is on *how* key people are identified as such, and what events take place to *initiate* their involvement in this regard. The relative visibility of key people within different role groups may depend on their relationship to those driving implementation. For example, if implementation of the EHR is driven primarily by clinicians and IT staff, the significance of administrators may not be immediately obvious to these project leads if their regular working practices do not expose them fully to the relevant functions of this group. Successful identification of key people and their *initiation* as such therefore requires detailed investigation of both formal and informal contributions within complex healthcare processes. Informal conversations with staff at all stages and levels of involvement can be just as valuable as more formal types of data (e.g. role descriptions) in informing both planning and evaluation of implementation with respect to *initiation.*

1.3.2 *How do participants become involved in the intervention? (Enrolment)*

Identifying participants and involving them in the work of implementation extends beyond key people; we also need to explore the practical processes by which others will be involved in implementing the intervention. Returning to the previous example (i.e. a new EHR within a GP clinic) we need to think about how different people will be bought in (or *enrolled*) as active participants. This is not the same as gaining consent to implement or change something but refers to the processes by which people become *actively* involved. Enrolment thus depends to a large degree on understanding the context in which participants operate, and again the focus is on *how* this occurs. For example, some initiatives may invite staff to take on specific tasks to drive implementation, and make them explicit points of contact for other staff affected by the intervention.

Implementation of the EHR might involve staff within different GP surgeries, or different role groups within the same surgery (e.g. reception, community nursing), actively seeking feedback on proposed changes and/or eliciting questions about implementation from colleagues.

1.3.3 *Why should a person participate? (Legitimation)*

Successful involvement of key people in the intervention, as well as wider enrolment of those working in the implementation space also implies that those involved believe that it is right for them to be, and that they can make valid contributions [3]. Organizational behavior change projects in health (such as improving infection control, or nutritional care for older people) often involve attempts to widen the sphere of concern with a particular activity (e.g. information governance, child protection, infection control) by, for example, stating that a given area is 'everyone's responsibility' [4–6]. Often, perceptions of legitimacy may be constrained by membership of specific professional groups (e.g. a nutritional care intervention might be seen initially as the exclusive responsibility of dietitians). Additional relational work is therefore often necessary to establish legitimacy with other groups. In the case of health informatics interventions, this may involve establishing relationships by meeting directly with clinicians and administrators using the system and establishing an understanding of *how* they will contribute to intervention and development.

1.3.4 *What processes will support people staying on task? (Activation)*

Projects in which participants have made sense of an intervention (*coherence*), identified key people (*initiation*), and bought those involved on board (*enrolment*) are well placed to begin initial implementation of their intervention. In these initial stages the tasks, relationships, and resources that have been established to support this work are *activated* – that is, they 'go live' and enter everyday work. These processes, being new, are vulnerable to various forms of disruption, particularly in settings where they compete with other tasks for the time and attention of participants. Processes associated with *activation* are the practical means by which those involved will be stay 'on the case', and how potential points of disruption may be identified and dealt with [3].

1.4 The operational work of implementation (Collective action)

Having made sense of the new set of practices and objects associated with the intervention (*Coherence*) and undertaken the relational work of understanding who should do what in the initial implementation of new practices (*Cognitive Participation*), we turn to the actual processes of implementation. Components in this construct highlight forms of operational work commonly necessary to support initial implementation.

1.4.1 *How does the intervention affect existing working practices and relationships? (Interactional Workability)*

Once the intervention goes live, is it in any way disruptive to normal ways of working? Does it 'get in the way' of other activities? While other constructs have pointed to sense making and organizational work in which key people may seek to anticipate these outcomes, it is often the case that complex interventions will require additional adaptive work as implementation progresses.

For example, a key problem of telemedicine systems historically has involved additional work required in communication and interpretation of complex clinical information, when compared with co-present consultations [7]. What we are interested in with respect to *interactional workability*, is the work that people have to do with objects (i.e. the physical implements that accompany an intervention, such as a new interface for patient record retrieval), new practices (e.g. a new way of performing diagnostic assessments), and each other to accommodate and adapt to new ways of working.

1.4.2 *How are confidence in, and accountability for the intervention built? (Relational Integration)*

Relational integration refers to forms of *knowledge work* that participants do to build accountability and maintain confidence in a set of practices and the people involved with them. Accountability can here be thought of as processes that give participants access to information (e.g. formal reports, or informal observations) about the outcomes of a given practice. Through such processes, confidence in an intervention and its associated practices and objects can be built and/or undermined. For example, confidence in a new teledermatology intervention was undermined when clinicians began to doubt the integrity of the images transmitted by the system, and began to examine patients in person alongside digitized images (resulting in greatly increased workload and increased pressure on their clinical department) [7]. Clinicians in this case undertook knowledge work that resulted in a loss of confidence in what was being transmitted, indicating not only *why* confidence was undermined, but *how*, and thereby identifying a point of failure at which such issues might be addressed (e.g. through development of image verification procedures that help clinicians to build accountability and confidence in the system).

1.4.3 *Who does what? (Skill set workability)*

Who should perform a given task? What are the processes for allocating responsibilities as the intervention progresses? Are they formal (for example, allocation by rota, or contractual changes to responsibilities), or informal through voluntary agreements between participants. Implementation of complex interventions often requires adaptation and renegotiation of roles and responsibilities, which can involve trade-offs between resource allocation (i.e. the time that specific people can contribute) and degree of need for specialist knowledge within a given part of the process. For example, a research group investigating the effectiveness of a decision aid for medication choice after a serious illness event had to decide whether the decision aid should be administered by trial managers with no clinical responsibility for the patient, or nurse practitioners actively involved in their care [3,8]. The trade-off here was between those with greater familiarity with and attachment to the intervention, compared with those closer to the field in which the decision aid intervened (i.e. the care pathway of patients recovering from serious illness events).

1.4.4 *Who gets what, and how? (Contextual integration)*

Successful interventions depend not only on individual and collective divisions of labour, but on allocation of resources to support them. *Contextual integration* looks at how practices and objects become (or fail to become) integrated within the wider context of the intervention setting, in terms of available resources. This component focuses on the processes through which resources are allocated as the intervention project progresses, involving questions such as: who has authority to allocate resources, and to what degree? Are those implementing the intervention able to access additional resources to deal with emergent challenges?

Returning to our earlier example of a GP surgery implementing a new EHR, we might explore whether administrators are able to access specialist knowledge support during implementation to help them work with the new system.

1.5 Evaluating implementation to promote embedding (Reflexive monitoring).

Having conducted the work of initial implementation, we need to consider how participants appraise the success of the implementation project as a whole, as well as the specific practices and objects associated with it. Components of this construct focus on kind of work done to evaluate the success of the intervention, analyse its impact on working relationships and individual practices, and (if necessary) make changes to it.

1.5.1 *How is information obtained to support appraisal work (Systemization).*

What informs how people appraise success, and how is this information obtained? Does information flow in the same way to all participants, or do some individuals and groups gather knowledge that others don't? Systemization may involve formal processes, such as the gathering of outcomes data within a randomized controlled trial. However, participants may also make use of information gathered through informal processes in both individual and collective appraisal (e.g. anecdotal examples of problems in practice) [3].

1.5.2 *How do participants work together to appraise the intervention? (Communal appraisal).*

Participants often work together to evaluate the worth of interventions (overall or in part), and these can involve formal processes such as team meetings, or informal groups (e.g. coffee break conversations). Different kinds of meeting may involve different processes that affect *how* the appraisal process is conducted. Formal meetings may have agendas structured around discussions of specific kinds of information (e.g. outcomes data from an RCT). They may also involve implicit or explicit divisions of labour that affect the kinds of information that enter discussions (e.g. formal meetings of consultants may exclude the informal observations of other participants). Likewise, unstructured appraisal may favour specific kinds of information (e.g. informal observations from practice) over others. The significance of this component is to recognize the kinds of appraisal work that are occurring within the field of implementation, and how these may affect the ways that participants understand and work with aspects of the intervention.

1.5.3 *How do participants evaluate the impact of the intervention individually? (Individual appraisal)*

Communal appraisal processes are also related to the work that individuals do to evaluate the impact of interventions of their own work, as well as the contexts in which they are set. Thinking back to our earlier example of the EHR within the GP surgery, individual clinicians may evaluate not only the worth of the programme, but its impact on their other tasks. If the system complicates and increases their workload, this may lower the value of the intervention to the clinician regardless of the overall impact on other areas of work within the surgery. From the point of view of understanding implementation, the focus here is on the processes by which individuals appraise the intervention, and the context in which different participants operate may influence this (i.e. the context and priorities of clinicians and administrators may differ relative to the other activities in which they are involved).

1.5.4 *Can participants modify aspects of the intervention, and if so how? (Reconfiguration)*

Appraisal work, both individual and collective, may lead to attempts by participants to modify practices associated with the intervention, or even aspects of the objects associated with it (e.g. diagnostic tools, patient information systems). For example, those leading implementation of the (hypothetical) EHR within a GP surgery may evaluate whether the benefits of the new system outweigh additional costs in terms of extra time or resource investment in using it. If they feel that aspects of the system negatively impinge on other important kinds of work, they may seek to modify aspects of the system. Depending on intervention design and the setup of implementation, this might involve asking developers to redesign some part of the front end, or to add features that allow it to integrate with other information systems within this space. However, more informal attempts may also be made to reconfigure how they work with the system ("work-arounds"), particularly if a route to requesting formal changes is not visible or practical. This may result in aspects of the new system being used alongside other systems or practices, in ways that were not anticipated by developers, and were not part of the original intervention design.

1.6 Relationships between the constructs

The ordering of constructs follows a general pattern from initial sense making, through organisational work to prepare for implementation, then the operational work of implementation, evaluation of its success, and potential reconfiguration. However, other kinds of connection between constructs are also possible, particularly following initial implementation when embedding the new procedure over a longer period may require revisiting or revising earlier steps. For example – work to set up the intervention, which would fall under the domain of cognitive participation, may reveal unforeseen implications for service, that require revisiting some earlier sense-making processes to address them and move forward. Similarly, evaluation work following initial implementation of an intervention that would map the *reflexive monitoring* construct, may reveal issues relating to how different people made sense of the intervention, that were not apparent until it went into service, and requiring changes to *coherence*-related activities. *Coherence*-related activities may also affect reflexive monitoring processes –

for example, if during *collective action* some feedback from a staff member indicates a significant area that the intervention is likely to affect, that was unlikely to be picked up by processes relating to reflexive monitoring, this could lead to changes in the programme design. Finally, we may simply see gradual changes made to the work of implementing interventions, which would come under collective action, because of findings related to reflexive monitoring. What these indicate is that implementation processes identified by the constructs may not proceed in a strictly linear fashion but encounter difficulties and go through revisions as they evolve.

1.7 From implementation to normalization – embedding new practices as routine aspects of care.

We began this section by noting that NPT focused on action, and we will conclude with an example indicating the importance of this for implementation in general. Figure 1 shows results from an investigation (a theory-led review) of systematic reviews into professional behaviour change in healthcare [9]. Along the top of this matrix we can see the NPT constructs and components, while the left-hand vertical distribution shows different types of professional intervention as defined by the Cochrane EPOC (Effective Practice and Organisation of Care) system. Studies in the review have been grouped by type, and each group has been ranked in terms of their success at effecting and sustaining professional behaviour change. The red boxes indicate which components of NPT were covered by the intervention, and these show a positive association between success of interventions, and their emphasis on the *collective action* and *reflexive monitoring* aspects of the intervention. On the basis of this, the review authors hypothesise that interventions which focus on attitudinal change are less likely to be effective in achieving long term behaviour change than those that reinforce new practice norms and associate them with peer and reference group behaviours [9]. Changing attitudes and building value are necessary activities, but may not suffice to ensure long-term success.

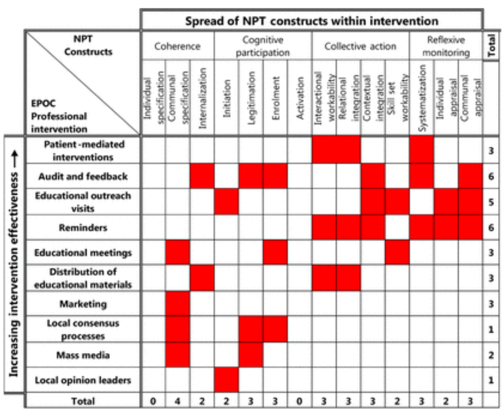


Figure 1 - Positive relationship between intervention effectiveness and focus on Collective action and Reflexive monitoring aspects of NPT [9].

2. Use of NPT in health informatics and service development

Having given an overview of NPT, we will now explore how the theory has been applied in both prospective planning, as well as ongoing and retrospective evaluation of implementations in health informatics contexts. We will examine two examples from published literature: firstly, an example of prospective planning of a digital healthcare intervention for management of diabetes [10,11]; and secondly an evaluation of implementation of a new electronic preoperative information system within a surgical pre-assessment clinic [12]. These two cases will serve as examples of how NPT has been used to plan and evaluate successful implementation of new health informatics systems, and identify mechanisms involved in this process.

2.1. Prospective planning of an evaluation study on implementation of a digital health intervention for Type-2 diabetes

Effective self-management is essential to good health outcomes and the prevention of associated complications for people with type 2 diabetes [10]. The UK National Institute for Heath and Clinical Excellence (NICE) recommends structured education to teach self-management; however, evidence suggests that only a small proportion of patients are offered this service, with fewer eventually attending [10]. Ross et al. developed an internet based self-management intervention: "HeLP-Diabetes: Healthy Living for People with Type 2 Diabetes", allowing patients to access self-management measures recorded by their GP surgeries, as well as information resources based on NICE guidance designed to complement existing in-person group education programme [10]. In planning implementation of *HeLP-Diabetes*, Ross et al. needed to consider how they would: determine uptake and use of the intervention by services and patients; identify factors promoting or restricting use; identify resources needed for successful implementation; and explore possible intervention effects on self-reported patient outcome measures [10]. The authors used NPT as an explanatory framework to explore the implementation process and guide interviews with NHS staff, using constructs and components as sensitizing resources (i.e. as indicators of general processes and kinds of work relevant to the outcomes of interest) [10]. Data collection also included informal feedback from staff at GP practices, collected by one researcher leading the implementation, as well as usage data from the *HeLP-Diabetes* software on number of patients signing up and the GP practices at which they were registered [11].

Ross et al. used NPT in analysis of interview, feedback, and usage data to develop an implementation plan for *HeLP-Diabetes*, in which specific implementation strategies were developed to target challenges mapped to the main constructs of NPT (see table 1). *Coherence*-related strategies included identifying key people within the local Clinical Commissioning Group (CCG – the body responsible for commissioning of services locally) as well as GP practice managers and leads [11]. This strategy allowed for targeted provision of educational materials emphasizing *HeLP-Diabetes* as an online system distinct from other self-management programmes, and its status as a free-to-use resource developed by a university. These strategies helped support *Differentiation* between *HeLP-Diabetes* and existing resources, and *Internalization* of value by drawing attention to its lack of cost to users, and the legitimacy of the developing body [11]. The implementation team also held educational outreach visits with healthcare professionals (HCPs) in which the nature of the programme, its evidence base, theoretical basis, participatory development, and benefits for patients as healthcare processes and organizations were discussed [11]. These conversations completed educational material by allowing potential adopters to explore questions regarding implications of the system for their own practice, as well as its efficacy, in deciding whether or not to adopt *HeLP-Diabetes* [11].

Table 1. Implementation strategies for *HeLP-Diabetes* targeting NPT constructs (adapted from [14]).

Coherence	Cognitive Participation	Collective Action	Reflexive Monitoring
Local opinion leaders	Interprofessional education	Educational meetings	Continuous quality improvement
Educational materials	Local consensus processes	Tailored interventions	Audit and feedback
Educational outreach visits	Educational materials		Reminders

Strategies to support *Cognitive Participation* included provision of a training session for HCPs to understand the actions and procedures necessary to ensure sustainable and successful implementation of the intervention [11]. Training sessions also included opportunities for staff within specific implementation sites to explore implementation with respect to local working contexts [11]. This was an important step in ensuring that implementation was flexible enough to accommodate planning for local contingencies (e.g. differences in how work is assigned within teams, methods of communication with patients).

Educational meetings and materials were also used to provide ongoing support for *Collective Action* processes during implementation. HCPs were given access to the *HeLP-Diabetes* system, allowing them to explore: how the intervention fitted the skill sets of staff; what resources might be necessary to support implementation at different sites; the knowledge necessary for HCPs to develop confidence in using the system; and how it might impact on interactions between colleagues, and with patients [11]. This process was supported by educational materials in the form of training booklets to support staff in becoming familiar with system functions (i.e. creation of a login, signing up a patient) [11].

Continual engagement with staff across the period of implementation also served a *Reflexive monitoring* function, as staff suggested that they would offer *HeLP-Diabetes* to patients more if they were receiving a greater number of related enquiries from patients [11]. This led to development of additional patient-focused advertising strategies to increase awareness including: TV screen adverts in waiting rooms; talks given at self-management groups; attendance at Diabetes UK events; coverage in practice newsletters; and mass mailouts to all patients at some implementation sites [11]. What is interesting to note here is the relationship between *Reflexive monitoring* in the form of staff feedback, and its use in revision of *Collective action* processes relating to *Interactional workability* (that is, a suggested change to the implementation strategy targeted at the relationship between HCP and patient) [11].

2.2. Evaluating implementation of a new electronic preoperative information system within a surgical pre-assessment clinic.

Surgical pre-assessment clinics (PACs) evaluate whether patients may be suitable for day case surgery or 23-hour care, or may require a longer in-patient stay. These have been introduced in Scotland as a result of policy recommendations intended to reduce unnecessary burden on services, and reduce surgical mortality rates [12]. PACs act as a gateway to surgical services from a wide range of referral pathways, involving multiple information flows. The PAC design evaluated by Bouamrane and Mair (2014) incorporated development and implementation of an electronic pre-operative information management system, to facilitate information sharing among members of the multidisciplinary PAC team. Development occurred iteratively by PAC staff in collaboration with the local NHS Health Board Information Technology team. In this article, the authors focus on one site (Dumfries and Galloway Royal Infirmary - DGRI) from a national study in Scotland. The authors modelled clinical processes using process-mapping techniques, and conducted 10 semi-structured interviews with five participants across four visits to the clinic site [12]. NPT was used in analysis of results from both process mapping and qualitative interview data.

The rationale for the clinic was found to be well established at DGRI, supported by previous institutional experience of problems with traditional in-patient routes lacking pre-assessment. In addition, the importance and relevance of the service was reinforced by national policy initiatives that incorporated performance targets. *Coherence* of the PAC in terms of overall relevance to strategic objectives of the institution was therefore well established. *Coherence* was also found to be high within the pre-assessment clinic, but less so at points of contact with other services. This was attributed in part to the number of different possible pathways to the PAC which were observed to be confusing to staff within the clinic, in addition to the fact that junior doctors involved in various routes to the PAC were not routinely involved in the clinic's assessment processes.

The collaborative design of PAC implementation and development, particularly with respect to the pre-operative information management system, was reported as a strength of the project. The authors report a 'teething period' of 12 months, after which specialist nurses leading PAC development were 'entirely satisfied' with information management practices. The combination of leadership from experienced pre-surgical nurses, and collaborative ongoing development with local NHS IT services ensured that key people relating to both clinical and health informatics aspects of the project were working together to drive forward development (an aspect of *Cognitive participation*).

Staff within the multi-disciplinary team (MDT) were found to be highly experienced in care and management of day-case patients. This foundation allowed PAC staff to effectively define their roles in relation to the new clinic, and build both individual and collective understandings of accountability (*collective action*). At the level of the wider institution however, participants expressed concerns with respect to replication by the PAC of information available through other sources (i.e. primary care). Here, the authors note that such concerns may in principle be addressed through improvements to integration of existing information systems. In the context of *Collective action*, such a development would require extending professional relationships through which roles are defined to encompass inter-departmental working (i.e. who is responsible for which tasks within an overall care pathway). Finally, although the prior experience of PAC staff was important in the success of the nurse-led clinic design process, there were no formal processes for continuing professional development or training at the PAC (*reflexive* *monitoring*). Transfer of knowledge on PAC procedures and related practice updates was observed to occur informally during other work, leaving the intervention vulnerable to staff attrition (as no formal process existed to ensure that this knowledge was transferred to new appointees).

2.3. Limitations of NPT

Before concluding with an exercise to help readers apply NPT to health informatics developments we will discuss some of its limitations, the first of which concerns lack of sensitivity to wider contextual factors beyond the immediate site of implementation. For example, Clarke et al. (2013) used NPT to evaluate implementation of a training programme for carers of stroke patients, within a cluster randomized controlled trial (RCT) [13]. The multi-site nature of this trial meant that variations in implementation context were present at the local level (e.g. service, resources, divisions of labour), in addition to regional and national policy changes (with differences in local responses to such changes providing further sources of complexity). In their evaluation, Clarke et al. noted that while NPT had been useful for identifying mechanisms and processes that inhibited implementation of the training programme, it did not capture the impact of these wider contextual factors [13]. At a national level, recruitment to the cluster RCT began shortly after the launch of a new National Stroke Strategy in 2008 [13]. In addition, many sites experienced competing demands on MDT members', patients', and care givers' time and resources from other service development initiatives [13]. All the hospital services involved were working towards the goal of stroke survivors spending all or part of their stay on a stroke unit, while most were also planning or introducing thrombolysis services. In addition, many sites were introducing early supported discharge schemes or reorganization of existing services, which required changes in staff locations and roles. While the impact of these factors may have been visible indirectly through their impact on other kinds of process identified through NPT (e.g. *Resource allocation*), Clarke et al. found that theory did not account fully for their role in the implementation context.

These observations indicate both the vulnerability of service developments (including health informatics innovations) to organizational turbulence, and how building relationships and processes that are resistant to such turbulence is essential in complex healthcare settings. They also indicate the importance of attending to contextual factors that shape implementation processes, a concern that has driven ongoing development of the theory [14]. In addition, authors such as Johnson et al. (2017) have sought to address these limitations in their application of the theory, by presenting adapted models that link the constructs with wider organizing structures and social norms (e.g. policies, public expectations of services, political contexts – see Figure 2). Elsewhere, in a systematic review of NPT use in feasibility studies and process evaluations, May et al. (2018) noted a number of additional criticisms from researchers: that NPT constructs overlapped; that the technical vocabulary of the theory was difficult; and that as a result coding qualitative data was difficult [1]. May et al. noted that problems of this nature seemed less evident when researchers used a more inductive approach to qualitative data analysis than they did when authors employed a framework approach [1].

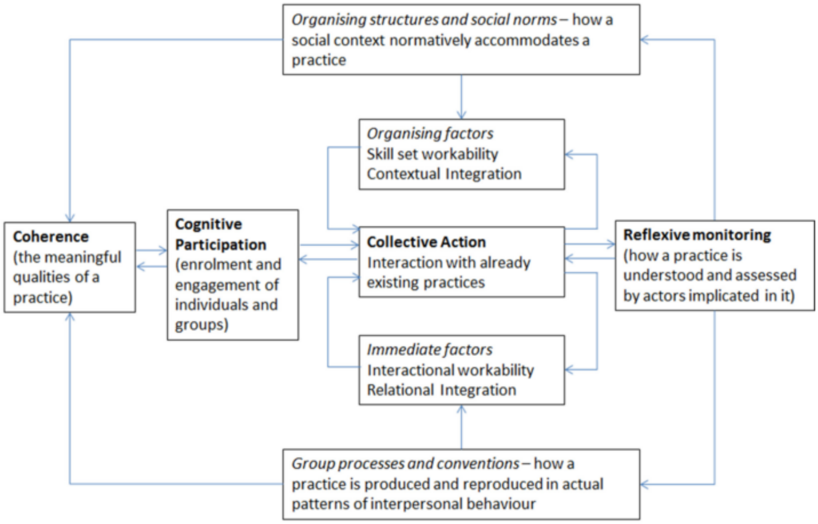


Figure 2. Johnson et al.'s (2017) adapted model of NPT [15].

3. Exercise

As we have seen, NPT can be used in prospective planning, as well as ongoing and/or retrospective evaluation of implementation of new technologies and processes. In this final section, we present a worked example of how NPT may be used to frame research questions for either purpose. It is important to note that this is only one way in which the theory has been used, and that other applications may be appropriate for different implementation projects. For an overview of how NPT has been applied in study design, data collection and analysis, we recommend that readers consult the 2018 systematic review conducted by May et al [1].

The exercise will involve using NPT to derive research questions in relation to implementation of a hypothetical health informatics system. To do this, we will use tables (see Appendices) containing descriptions of NPT constructs and components as a tool for linking research questions to components of the theory. We will provide a worked example for a single component, after which we invite readers to continue the exercise with remaining questions.

3.1. The scenario

Background - A community team of HCPs (comprizing nurses, occupational therapists, and dietitians) are implementing a new electronic patient record system for screening and treatment of malnutrition for patients in the community. These patients live in their own homes and are visited by members of the community team at regular intervals. Many are older, with multiple conditions including dementia, and as a result are at risk of undernutrition. Current team policy requires patients to be screened at monthly intervals using a clinical assessment tool, results of which are recorded and used to monitor nutritional health, and if necessary develop care plans for malnutrition.

Because of the distributed nature of visits, HCPs have hitherto recorded results of these screening assessments within a paper record. However, there is some concern that these paper records may not easily be integrated into team discussions of care planning (e.g. because the paper records are not always readily available). In order to address this, the team are now implementing an electronic system to record screening results, and retrieve them for team discussion and care planning.

Intervention: the team already use an electronic patient record system for most of the information regarding patient treatment and care, and the intervention updates this system to include a nutritional care component. Screening is performed on home visits by HCPs, who then enter the results into electronic record system. These records are then retrieved by team leaders and presented for discussion at team meetings, and where necessary care plans are agreed. These care plans are then entered into the system by team administrators who are present at the meetings and retrieved by HCPs prior to their next visit with the patient (care plans are integrated into patient information retrieval processes that already exist within the team). Outcome measures for success of the new system include numbers of patients screened (compared with previous years using the paper record), and changes in nutritional health of patients identified as being at risk of undernutrition.

Context: The community care team operates in a highly distributed fashion. Typically, HCPs will begin their shift by visiting the team base to retrieve patient records, after which they will begin their home visits. The work involves a range of patients with diverse needs and capacities, meaning that working conditions are variable and can be highly unpredictable, for example, a routine visit may uncover urgent care issues requiring immediate attention, reducing time for visits to other patients and increasing pressure on the individual HCP. Team meetings at which care planning takes place are also subject to time pressure. These may only last 30 minutes, during which 10 patients may be discussed, before HCPs are required to begin visits.

The intervention stems from concerns among managers and senior HCPs that this changeable working context often leads nutritional work to fall down the list of priorities, and that paper records of screening are vulnerable to exclusion from team care planning discussions because they aren’t stored in one place. In addition, embedding retrieval of nutritional care information within the existing patient records system was intended to reduce the time taken to source material for discussion, and reduce the likelihood that such information would be absent from care planning discussions.

Aim: The aim of this exercise is to identify questions that can be used to inform prospective planning and/or ongoing evaluation of implementation (readers are invited to explore one or both kinds of application depending on their interest). In both cases, the objective will be to identify factors that may promote or hinder implementation and longer-term embedding in routine practice, of the nutritional component of the electronic patient record system.

3.2. Generating research questions using NPT.

Construct	Construct description	Component	Topic of investigation	Planning questions	Evaluation questions
Coherence	*The sense-making work that people do individually and collectively when they are faced with the problem of operationalizing some set of practices.*	**Differentiation**	To understand how agents understand that a set of practices and their objects are different from each other.	How will information provided by the implementation team help HCPs distinguish the new procedure from current working practices?	Do community HCPs see the new procedure as different from existing ways of working, and if so how?
		Communal specification	Sense-making relies on people working together to build a shared understanding of the aims, objectives, and expected benefits of a set of practices		
		Individual specification	Sense-making has an individual component too. Here participants in coherence work need to do things that will help them understand their specific tasks and responsibilities around a set of practices.		
		Internalisation	Sense-making involves people in work that is about understanding the value, benefits and importance of a set of practices.		

Figure 3. Worked example use of NPT in nutrition screening and care planning scenario.

In Figure 3, we can see how a table might be used to identify questions relevant to the *Coherence* domain of NPT (see also Appendix). Note that wording of questions preserves the emphasis on *action*; for example, the planning question linked to the *Differentiation* component reads: '***How*** will information provided by the implementation team help HCPs distinguish the new procedure from current working practices?'. The 'how' is important here because while procedural differences between the two may seem obvious, it is possible that participants may interpret this process as a different way of doing the same thing (that is, they may not immediately see the benefits that those developing the intervention have in mind). Accounting for *how* these differences are made visible thus relates to an important part of the work necessary for successful implementation. For example, Ross et al's experience of implementing the *HeLP-Diabetes* intervention (discussed in section 2.1) indicates that discussions with HCPs, in addition to written information sources, were important in identifying and addressing questions about how the new procedure differed from existing practice [11]. Focus on action is also preserved in the example evaluation question (also linked to *Differentiation*): 'Do community HCPs see the new procedure as different from existing ways of working, and if so ***how***?' (see Figure 3). In both planning and evaluation, framing of the question will also affect further discussions about methods (i.e. a focus purely on belief or sentiment may suggest methods, such as attitudinal surveys using scale measures, that fail to capture processes relevant to understanding implementation which may better be investigated by, for example, interviews or in-person observations).

3.3. Next steps and concluding remarks

We now invite readers to continue with the example, by adding their own questions in the right-hand columns of the blank table (see Appendix). You may choose to do this for select components, or all of them[4] – or you may use the table to think through a different scenario of your own choosing.

Before concluding with some questions about method, it may also be helpful to note the use of this table in communicating the NPT framework to others involved in an implementation planning or evaluation project. Implementation projects can involve a range of professional groups, conventions, and languages which means that familiar examples may be helpful in building shared understanding of the general principles of NPT. Taking the example questions in the right-hand columns of figure 3, we can see how reading from the left-most column to this question links the component, construct, and context specific question. We can therefore also see that reading in the reverse direction offers an opportunity for communicating NPT principles using questions rooted in contexts that may be more familiar to some participants. This may be helpful in building a shared understanding of the framework – of implementing the implementation study itself.

Having derived research questions, the next step would be to consider research methods through which to conduct these investigations. Detailed discussion of the wide range of potential methods is beyond the scope of this chapter, and readers may look to the systematic review cited at the beginning of this section for a more detailed overview [1]. For those who may be implementing health informatics interventions, but be unfamiliar with process evaluations in general, this may provide a useful introduction to methods (particularly those involving qualitative observation) that have been used effectively in previous projects but may not feature commonly in other evaluations of health informatics systems.

References

[1] C.R. May, A. Cummings, M. Girling, M. Bracher, F.S. Mair, C.M. May, E. Murray, M. Myall, T. Rapley, and T. Finch, Using Normalization Process Theory in feasibility studies and process evaluations of complex healthcare interventions: a systematic review, *Implement. Sci.* **13** (2018) 80. doi:10.1186/s13012-018-0758-1.

[2] C. May, and T. Finch, Implementing, Embedding, and Integrating Practices: An Outline of Normalization Process Theory, *Sociology*. **43** (2009) 535–554. doi:10.1177/0038038509103208.

[3] C. May, T. Rapley, F.S. Mair, S. Treweek, E. Murray, L. Ballini, A. Macfarlane, M. Girling, and T.L. Finch, Normalization Process Theory On-line Users' Manual, Toolkit and NoMAD instrument, (2015). http://www.normalizationprocess.org.

[4] NHS Highland, Infection Control is everyone's responsibility, (2013). http://www.nhshighland.scot.nhs.uk/News/Pages/InfectionControliseveryone'sresponsibility.aspx (accessed November 18, 2018).

[5] Cheshire and Wirral Partnership NHS Foundation Trust, Safeguarding children is everyone's responsibility, 2017. http://www.cwp.nhs.uk/media/2040/safeguarding-children-is-everyoness-responsibility.pdf.

[6] Information Governance Alliance (IGA), Information Governance is the responsibility of everyone, 2015. https://digital.nhs.uk/data-and-information/looking-after-information/data-security-and-information-governance/information-governance-alliance-iga/information-governance-resources/new-resource-consultations/new-publications-closed-consultations/informatio.

[4] For further reference, readers may view a fully completed version of this table applied to an evaluation study of a nutritional care intervention (see Additional File 2 [17])

[7] T.L. Finch, F.S. Mair, and C.R. May, Teledermatology in the U.K.: lessons in service innovation, *Br. J. Dermatol.* **156** (2007) 521–527. doi:10.1111/j.1365-2133.2006.07608.x.
[8] T. Karagiannis, A. Liakos, M.E. Branda, E. Athanasiadou, M. Mainou, P. Boura, D.G. Goulis, A. LeBlanc, V.M. Montori, and A. Tsapas, Use of the Diabetes Medication Choice Decision Aid in patients with type 2 diabetes in Greece: A cluster randomised trial, *BMJ Open*. **6** (2016) 1–10. doi:10.1136/bmjopen-2016-012185.
[9] M.J. Johnson, and C.R. May, Promoting professional behaviour change in healthcare: What interventions work, and why? A theory-led overview of systematic reviews, *BMJ Open*. **5** (2015). doi:10.1136/bmjopen-2015-008592.
[10] J. Ross, F. Stevenson, C. Dack, K. Pal, C. May, S. Michie, S. Parrott, and E. Murray, Evaluating the implementation of HeLP-Diabetes within NHS services: Study protocol, *BMC Health Serv. Res.* **14** (2014) 1–8. doi:10.1186/1472-6963-14-51.
[11] J. Ross, F. Stevenson, C. Dack, K. Pal, C. May, S. Michie, M. Barnard, and E. Murray, Developing an implementation strategy for a digital health intervention: an example in routine healthcare, *BMC Health Serv. Res.* **18** (2018) 794. doi:10.1186/s12913-018-3615-7.
[12] M.-M. Bouamrane, and F.S. Mair, A study of clinical and information management processes in the surgical pre-assessment clinic, *BMC Med. Inform. Decis. Mak.* **14** (2014) 22. doi:10.1186/1472-6947-14-22.
[13] D.J. Clarke, M. Godfrey, R. Hawkins, E. Sadler, G. Harding, A. Forster, C. McKevitt, J. Dickerson, and A. Farrin, Implementing a training intervention to support caregivers after stroke: A process evaluation examining the initiation and embedding of programme change, *Implement. Sci.* **8** (2013) 96. doi:10.1186/1748-5908-8-96.
[14] C. May, Towards a general theory of implementation, *Implement. Sci.* **8** (2013) 18. doi:10.1186/1748-5908-8-18.
[15] M.J. Johnson, A.A. Leaf, and F. Pearson, Successfully implementing and embedding guidelines to improve the nutrition and growth of preterm infants in neonatal intensive care: a prospective interventional study, *BMJ Open*. (2017). doi:doi:10.1136/bmjopen-2017-017727.
[16] J.C. Moullin, D. Sabater-Hernández, F. Fernandez-Llimos, and S.I. Benrimoj, A systematic review of implementation frameworks of innovations in healthcare and resulting generic implementation framework, *Heal. Res. Policy Syst.* **13** (2015) 1–11. doi:10.1186/s12961-015-0005-z.
[17] M. Bracher, J. Murphy, K. Steward, K. Wallis, and C.R. May, What factors promote or inhibit implementation of a new procedure for screening and treatment of malnutrition in community settings? A prospective process evaluation of the *Implementing Nutrition Screening in Community Care for Older People* (INSCCO, *BMJ Open*. **9** (2019) e023362. doi:10.1136/bmjopen-2018-023362.

Appendix – Identification tables for questions relating to planning and/or evaluation of a complex healthcare intervention using NPT.

Construct	Construct description	Component	Topic of investigation	Planning questions	Evaluation questions
Coherence	*The sense-making work that people do individually and collectively when they are faced with the problem of operationalizing some set of practices.*	**Differentiation**	To understand how agents understand that a set of practices and their objects are different from each other.		
		Communal specification	Sense-making relies on people working together to build a shared understanding of the aims, objectives, and expected benefits of a set of practices		
		Individual specification	Sense-making has an individual component too. Here participants in coherence work need to do things that will help them understand their specific tasks and responsibilities around a set of practices.		
		Internalisation	Sense-making involves people in work that is about understanding the value, benefits and importance of a set of practices.		
Cognitive Participation	*Cognitive Participation is the relational work that people do to build and sustain a community of practice around a new technology or complex intervention.*	**Initiation**	When a set of practices is new or modified, a core problem is whether or not key participants are working to drive them forward.		
		Enrolment	Participants may need to organize or reorganize themselves and others in order to collectively contribute to the work involved in new practices. This is complex work that may involve rethinking individual and group relationships between people and things.		
		Legitimation	An important component of relational work around participation is the work of ensuring that other participants believe it is right for them to be involved, and that they can make a valid contribution to it.		
		Activation	Once it is underway, participants need to collectively define the actions and procedures needed to sustain a practice and to stay involved.		

Construct	Construct description	Component	Topic of investigation	Planning questions	Evaluation questions
Collective Action	*Collective Action is the operational work that people do to enact a set of practices, whether these represent a new technology or complex healthcare intervention.*	**Interactional Workability**	This refers to the interactional work that people do with each other, with artefacts, and with other elements of a set of practices, when they seek to operationalize them in everyday settings.		
		Relational Integration	Knowledge work that people do to build accountability and maintain confidence in a set of practices and in each other as they use them.		
		Skill set Workability	Allocation work that underpins the division of labour that is built up around a set of practices as they are operationalized in the real world.		
		Contextual Integration	Resource work - managing a set of practices through the allocation of different kinds of resources and the execution of protocols, policies and procedures.		
Reflexive Monitoring	*Reflexive Monitoring is the appraisal work that people do to assess and understand the ways that a new set of practices affect them and others around them.*	**Systematisation**	Participants in any set of practices may seek to determine how effective and useful it is for them and for others, and this involves the work of collecting information in a variety of ways.		
		Communal appraisal	Participants work together - sometimes in formal collaboratives, sometimes in informal groups to evaluate the worth of a set of practices. They may use many different means to do this drawing on a variety of experiential and systematized information.		
		Individual appraisal	Participants in a new set of practices also work experientially as individuals to appraise its effects on them and the contexts in which they are set. From this work stem actions through which individuals express their personal relationships to new technologies or complex interventions.		
		Reconfiguration	Appraisal work by individuals or groups may lead to attempts to redefine procedures or modify practices - and even to change the shape of a new technology itself.		

Part 3

Synthesis

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

doi:10.3233/SHTI190123

The NASSS Framework – A Synthesis of Multiple Theories of Technology Implementation

Trisha GREENHALGH[a,1] and Seye ABIMBOLA[b]
[a]*Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK*
[b]*School of Public Health, University of Sydney, Sydney, Australia, The George Institute for Global Health, Sydney, Australia*

Abstract. Technologies are often viewed as the route to better, safer and more efficient care, but technology projects rarely deliver all the benefits expected of them. Based on a literature review and empirical case studies, we developed a framework (NASSS) for studying the non-adoption, abandonment and challenges to scale-up, spread and sustainability of technology-supported change efforts in health and social care. Such projects meet problems usually because they are too complex – and because the complexity is sub-optimally handled. NASSS consists of six domains – the illness or condition, the technology, the value proposition, the individuals intended to adopt the technology, the organisation(s) and the wider system – along with a seventh domain that considers how all these evolve over time. The NASSS framework incorporates a number of other theories and analytic approaches described elsewhere in this book. It is not intended to offer a predictive or formulaic solution to technology adoption. Rather, NASSS should be used to generate a rich and situated narrative of the multiple influences on a complex project; to identify parts of the project where complexity might be reduced; and to consider how individuals and organisations might be supported to handle the remaining complexities better.

Keywords. NASSS framework; complexity of innovations; diffusion of innovation; value proposition; scale-up

Learning objectives

After reading this chapter the reader will be able to:

1. Articulate various individual theories of technology adoption and implementation within a multi-level integrated framework.
2. Draw different theories together to explain the multiple and complex challenges to the adoption, scale-up, spread and sustainability of technology-supported programmes in healthcare.
3. Design an evaluation of a health informatics intervention based on the NASSS framework.

[1] Corresponding Author, Professor Trisha Greenhalgh, E-mail: trish.greenhalgh@phc.ox.ac.uk.

1. Introduction to the NASSS framework

1.1. Origins and overview of the NASSS framework

Most research into technological innovations has focused on technology development and mapping patterns of adoption, with little attention paid to the systematic study of the *non-adoption* of promising technologies. This chapter introduces an evidence-based framework (abbreviated NASSS) for studying the **n**on-adoption and **a**bandonment of technologies by individuals and the challenges to **s**cale-up, **s**pread and **s**ustainability of such technologies in health and care organizations. The NASSS framework was developed using two parallel processes: a narrative systematic review of theory-informed frameworks for analysing and evaluating technology-supported change programs in health and social care [1], and empirical testing and iterative refining of the NASSS domains using a diverse sample of technology implementation projects, written up as rich mixed-method case studies followed up for (at the time of writing) three years [2].

The NASSS framework is shown in Figure 1. It consists of seven domains, each of which may be simple (few components, predictable), complicated (many components but still largely predictable) or complex (many components interacting in a dynamic and unpredictable way). The more complexity there is in the system, the less likely the technology is to achieve sustained adoption across the system (and the more likely it is to be abandoned). The different sub-domains in the NASSS framework (right-hand panel in Figure 1) can be applied adaptively to produce a nuanced narrative that reveals the different kinds of complexity in the unfolding programme.

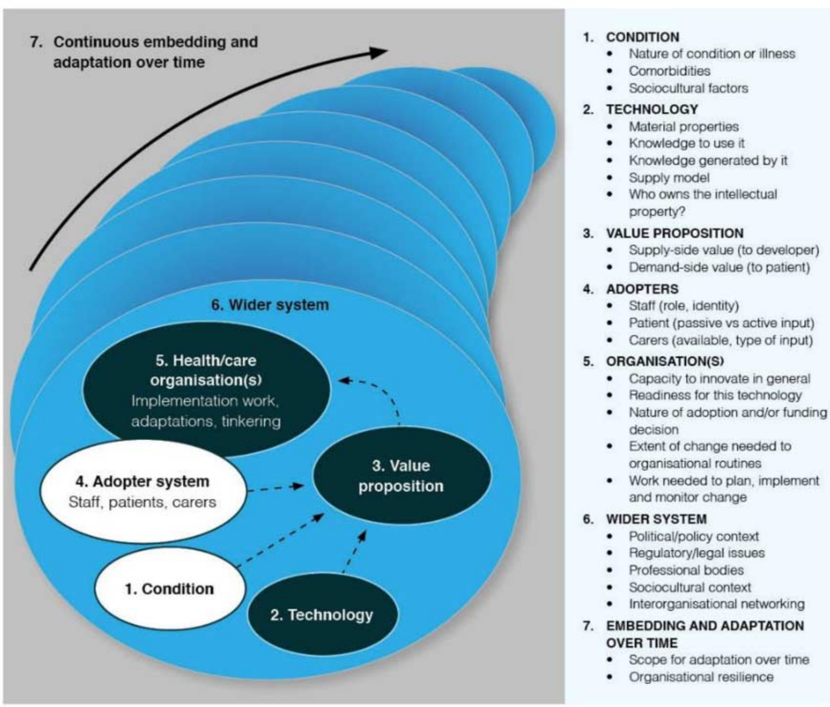


Figure 1: The NASSS framework for studying non-adoption and abandonment of technologies by individuals and the challenges to scale-up, spread and sustainability of such technologies in health and care organisations (adapted from Greenhalgh et al [1])

The NASSS framework is not a theory on its own. It is a map of possible areas of complexity to take into account when planning, analysing or writing up a project or initiative involving a technology. In addition to an over-arching theory of system complexity (which can be thought of as a ‘grand theory’ – that is, one at a very high level of abstraction and generality), each domain in the NASSS framework may be informed by one or more focused (‘middle-range’) theories, many of which are explained in more detail elsewhere in this book. Below, we introduce the over-arching theory of complex adaptive systems which informs the NASSS framework as a whole, followed by examples of relevant underpinning theory(ies) for each individual NASSS domain.

1.2. The importance of complexity

Complexity has been defined by Cohn et al as “*a dynamic and constantly emerging set of processes and objects that not only interact with each other, but come to be defined by those interactions*” (page 40) [3]. Complex [adaptive] systems are characterised by fuzzy boundaries; their interacting agents operate according to internal rules that cannot always be predicted; such systems interact, adapt and co-evolve with other systems [1, 4, 5]. Whilst it is fashionable in healthcare circles to talk about complex interventions, it is important to recognise that complexity is a feature not just of an intervention but of the system(s) into which the intervention is introduced [6, 7]. Indeed, even when an intervention (such as a technology) is simple (defined as having one active component and unchanging) rather than complex (defined as having multiple interacting components and perhaps also evolving over time), the *system* will almost invariably need to adapt in some way to accommodate it [6, 7]. Typically, a planned technological intervention (such as a patient-facing portal to access a health provider) and its context (e.g. a deprived rural community with unreliable broadband coverage) will be inter-related and reciprocally interacting.

Complex systems have many other features that are relevant to the study of technologies in a health care context. It is simply not possible to predict with certainty what will be the outputs if X is the input. Health systems are rapidly changing (the baseline against which the implementation is being evaluated is rarely static). Technologies may be more or less reliable in different contexts (software, as we all know, has a tendency to crash or develop bugs when interfaced with other software). Work-as-imagined (the guideline or standard operating procedure) necessarily differs from work-as-done [4]. Human actions may be variously constrained (both materially and socio-culturally).

In complex systems, therefore, decisions must often be made on the basis of incomplete, contested or only partially relevant data. Furthermore, certainty not only eludes us *now*, it will continue to elude us as the project progresses and we will have to learn to work with *uncertainty*. Indeed, the conclusion “more research is needed” often needs to be replaced with “more pragmatism is needed”. In such systems, human agents use their creativity and generate pragmatic solutions that make sense locally – at least for a while, until circumstances change, when they must adapt again. When researching complex systems, we need to surface and celebrate (rather than ignore or sanitise) all the articulations, workarounds, muddling-through and emergent activity that keep the show on the road.

A complex systems perspective holds that the planning, analysis and writing up of a technology project should be much more than a linear account of a particular goal and the extent to which it was met. It should be a richly-described case study (which may contain quantitative as well as qualitative data) of how human actors made it happen *despite* all the uncertainties, contingencies, inconsistencies, material challenges and micropolitical hiccups – and how the goals changed (perhaps quite appropriately) as the project unfolded and contextual influences changed [8].

1.3. Domains of the NASSS framework

Against this background of complexity in health systems, let us now consider the different domains of the NASSS framework, shown in Figure 1, and the different kinds of complexity that can occur. Broadly speaking, such complexity can be logistical (relating to the scale, scope and different inter-related sub-systems involved) or socio-political (relating to personal, interpersonal or inter-organisational issues such as differences in values or conflicts of interest).

Domain 1 in the NASSS framework is the condition (perhaps an illness, such as diabetes, or risk state, such as increased tendency to falls). The human body is of course a complex system, as is the family and community in which the sick person is cared for. The most obvious theoretical influences on this domain are biomedical and epidemiological theories of disease (which often but not always allow prediction of how the condition and its co-morbidities will progress over time) and pharmacological theories of how drugs work and interact. In addition, a number of theories of *illness* (that is, disease as experienced by the patient) are relevant here. Sociological framings depict illness as a unique personal (and family) experience which may involve stigma, biographical disruption, loss of status, reduced income and a heroic struggle to retain dignity, rebuild identity and live a moral life in the face of adversity [9, 10]. Political economy framings depict illness as the result of poverty or maldistribution of power in society (for example, Julian Tudor Hart's Inverse Care Law states that people most in need of health care are least likely to seek it or receive it) [11].

Complexity in Domain 1 may occur, for example, when the condition is metabolically volatile (e.g. sepsis), inherently unstable (e.g. alcohol dependency), poorly described or understood (e.g. a newly described syndrome), associated with multiple co-morbidities and polypharmacy (for example, in older people) or influenced by socio-economic or cultural factors (including poverty and material circumstances; limited access to healthcare; low health literacy, system literacy or digital literacy; cultural traditions and norms; social exclusion). For an overview of the kinds of complexity that affect the condition or illness, see this review [12].

Domain 2 is the technology, for which a number of underpinning theories covered in separate chapters elsewhere in this text book may be relevant, including socio-technical systems theories[2], technology adoption theories[3], normalisation process theory[4] and user-centred design theories[5]. In our own empirical work applying NASSS to patient-facing technologies (e.g. designed to support self-care in the home), we have drawn particularly on Jeanette Pols' theory interpretation of actor-network theory, which

---

[2] See Chapter 7, "Distributed Cognition: understanding complex sociotechnical informatics" and Chapter 8, "Using Actor-Network Theory to study health information technology interventions".

[3] See Chapter 6, "Technology Acceptance Models in health informatics: TAM and UTAUT".

[4] See Chapter 15, "Implementing and embedding health informatics systems – understanding organisational behaviour change using Normalization Process Theory (NPT)".

[5] See Chapter 5, "Linking Activity Theory with User Centred Design: a human computer interaction framework for the design and evaluation of mHealth interventions".

focuses on how particular technologies bring particular kinds of knowledge into play and render other kinds of knowledge less visible [13].

Complexity in Domain 2, therefore, may relate to the material properties and functionality of the technology itself (especially its dependability and speed of operation); to the knowledge needed to use it (hence, to how easily staff and patients can be trained); or to the knowledge it brings into play and how much that knowledge is likely to be trusted or contested. It may also relate to the technology supply model (e.g. to what extent is the technology substitutable?) and to the intellectual property (IP) it generates (how easy is it to say who 'owns' the IP?).

Domain 3 is the value proposition – both supply-side (value to the developer and/or healthcare system) and demand-side (value to the patient and/or insurer). Relevant to Domain 3 are various theories of value generation. Here, we describe one: transaction costs theory, which was developed to explain the governance implications of costs that constitute friction or a barrier to otherwise desirable economic or social exchange [14, 15]. The level of transaction costs on any given patient-provider interface is influenced by technology; and can be measured in terms of the number and duration of steps involved in patient pathways and clinical workflows [16, 17]. Transaction costs include search and information costs, bargaining and payment costs, or monitoring and enforcement costs. From the perspective of a consumer transaction costs are all costs incurred by the consumer that are not transferred to the seller (e.g. the time spent obtaining information on the good or service, and on prices and potential alternatives, legal fees, and the costs of establishing credibility as a buyer). From the perspective of a producer, transaction costs are all costs which the producer would not incur were they selling the good to themselves (e.g. time spent waiting while people examine the good or service, agent and advertising fees, and the costs of establishing credibility as a seller) [18].

In the software-platform revolution that began 20 years ago with the launch of eBay, and continues with Uber and Airbnb, Munger argues that "entrepreneurs have for the first time been able to specialise in selling not more stuff, but reductions in transaction costs for access to existing stuff" (page x) [19]. These platforms reduce transaction costs by providing "(1) information about identity and location [of potential transacting partners]; (2) a way of making payment that both parties can trust; and (3) a way of outsourcing trust on performance of the terms of the contract" (page 393) [19]. Often, the primary value of a technology is reduction in transaction costs; reducing frictions on the patient-provider interface – for example, patient-facing digital health innovations such as video consultations, and apps designed to support self-care in the home or help patients to locate, pay for, and rate laboratory services in their vicinity. But by reducing transaction costs, a technology may also add value by creating/maximising the capacity of a system to deliver services (on the supply side) or by creating/maximising opportunities for population access to services (on the demand side).

Complexity in Domain 3 thus relates to difficulties in formulating a plausible business case for developing the technology or to verifying the assumptions about how value will be generated [20, 21]. A simple value proposition offers a clear business case for investors *and* evidence that patients and the health service will benefit. In a complex situation, the business case for developing the product is implausible, or rests on unverifiable assumptions, and/or the results of health technology assessment studies are unavailable or contested. In addition, a business case may be complex when it is unclear or unpredictable how the innovation will re-distribute transaction costs among stakeholders, and how this may change over time – for example, video consultations may reduce transaction costs for patients (e.g. direct and indirect costs of travel) but increase transaction costs for the health system (e.g. costs of installing the videoconferencing equipment); conversely, a self-monitoring app may increase transaction costs for patients (e.g. forgone income due to time spent on monitoring by the patient or their family) but reduce transaction costs for the health system (e.g. less expense on staff time for patient monitoring).

Domain 4 is the adopter system: the staff, patients and carers who will be expected to use the technology (but who may refuse to use it or find they are unable to use it). Relevant theories here include theories of how people learn to use technology (one example is Bandura's social learning theory, which emphasises on-the-job learning and the importance of respected role models [22]). But non-use of a technology is rarely solely due to lack of knowledge or skill. We may also need to invoke sociological theories of why professionals resist new technologies, (see for example Greenhalgh, Stones and Swinglehurst's adaptation of Giddens' structuration theory to explore professional resistance to nationally mandated software programmes [23]). Complexity in Domain 4 occurs not only when using the technology requires knowledge or skills the user does not have but also when the roles and practices assumed by the technology threaten deeply held values or norms – for example, when a staff member is expected to do something she feels is against her professional code of conduct or work in a way that provides what she feels is a lower standard of care.

In relation to adoption of technologies by patients, May's burden of treatment theory (like transaction costs theory) proposes that shifting the work of care from clinician to patient places new demands on the sick, hence may be disempowering rather than empowering [24]. Such work may include taking readings and entering data (e.g. in many telehealth applications), making judgements (e.g. about what is an emergency or whom to contact in a crisis) or adjusting medication (for example, in response to a treatment titration algorithm).

Domain 5 is the healthcare organisation(s). The theoretical underpinning of this domain was summarised in an earlier paper, 'Diffusion of Innovations in Health Service Organizations' [25], which included an extensive systematic review of the characteristics of organisations that support innovation. These included theories of organisational *structure and climate* (for example, the well-documented findings that well-led organisations with flat hierarchies, devolved decision-making, slack resources and a risk-taking climate find it easier to innovate than those lacking these features), theories of *absorptive capacity* (preconditions for capturing knowledge from outside the organisation and disseminating it internally), theories of organisational *readiness* (especially the notion of innovation-system fit and the potential 'wrecking power' of strategically-placed opponents) and various theories of assimilation and implementation. In addition, theories of incremental versus disruptive change are relevant ('disruptive innovation' of the rip-and-replace school succeeds far less often than a more incremental approach to change [26]). Finally, May's normalisation process theory[6] unpacks the work of implementing a technology in an organisation, including *coherence work* (the work that people do to make sense of a practice), *cognitive participation* (work to enrol and engage other people in relation to that practice), *collective action* (work to enact the new practice), and *reflexive monitoring* (the work involved in evaluating the impact of the technology) [27].

[6] See Chapter 15, "Implementing and embedding health informatics systems – understanding organisational behaviour change using Normalization Process Theory (NPT)".

Complexity in Domain 5, then, may relate to the organisation's general capacity to innovate (such as leadership, clinician-managerial relationships, absorptive capacity for new knowledge and availability of slack resources); its readiness for this particular technology (tension for change, balance of supporters and opponents); the nature of the adoption and funding decision (more complex if it depends on inter-organisational agreements and speculative cross-system savings); potential disruption to existing routines (the less there is, the simpler it will be); or the extent of work needed to implement the changes (including ensuring staff buy-in, delivering the change and evaluating the change).

Domain 6 is the wider system. There are many potentially relevant theories that suggest how external social, political, technological and economic context may affect the uptake of innovations. One example is Richard Scott's neo-institutional theory, which proposes that innovation and change in healthcare organisations is heavily influenced (and may be slowed down) by three broad types of social forces or "institutional pillars": regulative (laws, regulations and contracts which stipulate what *must* happen), normative (professional and societal expectations about what *should* happen) and cultural-cognitive (taken-for-granted scripts and mental models about what generally *does* happen). Each pillar offers a different rationale for legitimising human action or inaction, by virtue of being (respectively) legally sanctioned, morally (e.g. professionally) authorised, or culturally supported. The wider system also embraces the networks that exist between organisations and theories of how networking and knowledge-sharing between organisations can significantly increase the uptake and embedding of innovations within them [25].

Complexity in Domain 6 may relate to negative perceptions of the innovation or specific blocks to its introduction from policymakers, regulatory or professional bodies, or the general public [8]. It may also indicate limited scope for networking activities among organisations (for example via quality improvement collaboratives), which are known to improve organisations' capacity to innovative.

Domain 7 is continuous embedding and adaptation over time (of both the technology and the service or organisation). Relevant theory here includes Everett Rogers' consistent finding that 'potential for reinvention' is a key determinant of successful adoption of an innovation [28], and also to the notion of organisational resilience [29], which has been defined as *"the intrinsic ability of a system to adjust its functioning prior to, during, or following changes and disturbances so that it can sustain required operations, even after a major mishap or in the presence of continued stress"* (page 1) [30]. Complexity in Domain 7 may thus relate to the technology's lack of potential to adapt to changing context or to the organization's lack of resilience.

2. Usage of NASSS framework in health informatics: A case study of a telehealth system for heart failure *(SUPPORT-HF)*

Empirical studies by our own team [2, 31] and others (as yet unpublished) have demonstrated the value of the NASSS framework for constructing a rich narrative of an unfolding technology-supported change programme and identifying the various interacting uncertainties and interdependencies that need to be contained and managed if the programme is to succeed.

An example from our empirical dataset is a home-based telehealth system for heart failure, known as SUPPORT-HF, which provided remote data on patients' blood pressure, blood oxygen levels and heart rate and rhythm [32]. When aggregated over time, these data could alert clinicians to impending deterioration, prompting a phone call, an invitation to clinic or a home visit.

The technology was adopted patchily (even by participating sites in a randomised controlled trial, all of whom had initially agreed to participate), and the service model which it supported was not straightforward enough to implement in community or hospital-based heart failure clinics. Below, we consider the different kinds of complexity in the SUPPORT-HF study and how the different NASSS domains can help in analysing this complex case.

The condition: Heart failure affects 1–4% of the adult population; it is commoner in ethnic minorities and people from socio-economically poor backgrounds, and its prevalence increases with age (the average age of first diagnosis of heart failure patient is 76) [33]. It has multiple causes and complex pathophysiology; heart failure that results from an isolated defect in a part of the heart (e.g. a leaky valve) is now much less common than heart failure linked to general deconditioning in an obese person who also has high blood pressure and diabetes. On average, four to five comorbidities add to symptom and treatment burden and influence prognosis. Co-existing frailty, depression and cognitive impairment are common. The course of the condition is highly variable but it can lead to rapid deterioration and/or sudden death. Heart failure frequently causes extreme fatigue and may cause confusion; patients typically describe themselves as bewildered and frightened.

The technology: The SUPPORT-HF technology consisted of standardised instruments for biomarker monitoring (weight, blood pressure, heart rate) along with a tablet computer (which had been developed using a co-design methodology) into which patients entered data for remote transfer to a monitoring centre. Participants in both arms of the trial received the technology and automated feedback messages (e.g. if results went outside pre-set parameters). In the intervention arm, the patient's family physician was alerted to out-of-range results and offered suggestions for changes in therapy, whereas in the control arm, results were made available on a Web portal for the patient's physician to access if they chose to.

The value proposition: Because the technology had been developed as a research initiative, the value chain was somewhat speculative. The assumption was that using telehealth would enable the hard-pressed community heart failure nurses to take on a higher case load (from 35-50 patients per nurse to an estimated 200 patients per nurse). This was thought to be possible because of reduced travel time for nurses (who did a lot of their work by home visiting) and the assumption that processing remote data (blood pressure, heart rate and rhythm, body weight) would be a quicker way to monitor the course of a patient's heart failure than undertaking regular clinical examinations of the patient. However, for various reasons, the trial was slow to recruit (in some but not all sites) and many patients were either not entered into the study (because the care package was considered clinically inappropriate) or because they were unable or unwilling to undertake the monitoring, or because broadband was unavailable. Thus, at the time of writing, the anticipated economies of scale in relation to nurse caseload have not yet been realised (and may have been over-optimistic). Another as-yet unknown transaction cost of the telehealth model is the cost of supporting and maintaining the technology in patients' homes.

The intended adopters: Staff at the different SUPPORT-HF sites engaged variably with the study, sometimes leading to slower than predicted recruitment. Some heart failure nurses were extremely keen but others engaged only superficially with the trial protocol and recruited few patients, citing previous poor experiences with telehealth, concern that a remote monitoring service would threaten their jobs, or a belief that patients 'deserved better'. The implied role change for the specialist heart failure nurse was potentially far-reaching. Instead of spending her time seeing patients in clinic or visiting them at home, nurses would now be spending a proportion of their time sitting in a data processing centre looking at on-screen data and trends. Furthermore, one driver for the introduction of the telehealth programme was a rapidly rising incidence of heart failure (and, because of improved care, patients were surviving many years after diagnosis). One cardiologist spoke of a health economic model in which the case load for each nurse would increase from 35 to 200 patients. Whilst some nurses embraced this vision enthusiastically, others strongly resisted it on the grounds that a dramatic reduction in direct patient-facing activity meant that they were no longer *being* heart failure nurses.

Patients expressed a wide range of views about remote biomarker monitoring in the SUPPORT-HF study; some took an active interest in their readings, engaged enthusiastically with the feedback they received, and found this monitoring reassuring. Others found the experience confusing and burdensome; they did not know (and did not wish to learn) what the numbers meant. In some cases, a research nurse who knew the patients well provided (unofficial) telephone support to maintain engagement.

The organisations: Participating sites in the SUPPORT-HF study were generally semi-autonomous cardiology units based in large district general or teaching hospitals. With few exceptions, leadership and managerial relations were good and (because of research support funding for the trial) there was sufficient financial slack to support introduction of the technology. As a research initiative, the SUPPORT-HF technology was not integrated into mainstream services, but we tentatively predict that because of the major knock-on implications for work routines (especially in relation to community heart failure nurses), this technology will be experienced as 'disruptive' and hence prove difficult to mainstream after the 'proof of concept' phase ends. One further external factor is the complexity of heart failure services, which typically span general practice, community clinics and hospital services – each of which has a different funding stream and different patient caseload. A telehealth-supported service in one of these sectors may need to interface with other sectors in the same locality that do not support (and perhaps do not trust) telehealth.

The wider system: The SUPPORT-HF study unfolded at a time when there was a strong policy push for telehealth initiatives in general and for initiatives to reduce outpatient attendance in particular. But whilst the policy environment was positive, our data showed that in some sites up to half the eligible patients could not be randomised because of the variability of broadband speed outside the main cities. The extent of inter-organisational networking among participating departments in the SUPPORT-HF study was limited as this was not an explicit component of the trial intervention; we suggest that if this technology is introduced as a business-as-usual intervention post-trial, networking and knowledge-sharing among organisations should be supported (either via a virtual link or occasional face to face meetings).

Evolution and adaptation over time: The tablet technology used for SUPPORT-HF included some limited scope for adaptation and customisation, but our qualitative data suggested that both staff and patients wished to adapt it further (either to accommodate individual needs and preferences or to adjust to external factors such as a changing technical infrastructure in the participating service). We are somewhat pessimistic about this particular technology's potential for surviving into the future, but the same service could be delivered on a substituted technology.

In summary, the main complexities in the SUPPORT-HF example are the condition itself (heart failure is serious, unpredictable, heterogeneous, associated with multiple comorbidities and occurs more commonly in patients who are poor and from minority ethnic groups), untested assumptions in the value proposition (such as predicted uptake and the cost of processing remote data), the intended adopters (neither staff nor patients view the technology with unqualified enthusiasm, and a key staff group may perceive a threat to their scope of practice and job security), and the disruptive implications of the technology for organisational (and especially inter-organisational) routines. Furthermore, lack of broadband access in rural and remote parts of the UK currently preclude this technological model as a solution in the very geographical regions where it could potentially be most useful.

3. Discussion

The NASSS framework has been developed relatively recently; whilst many teams around the world are currently exploring its potential, published studies of its application are limited. Indeed, we are still at the stage of formulating hypotheses which we encourage others to test. At the most broad-brush level, for example, we hypothesise that:

- when most or all of the NASSS domains can be classified as *simple*, the programme is likely to be easy to implement and to be achieved on time and within budget;
- when many domains are classified as *complicated*, the programme will be achievable but it will be difficult and likely exceed its timescale and budget;
- when multiple domains are *complex*, the chances of the programme succeeding at all are limited.

The reality is that almost no technology projects in health and social care are simple. Therefore, to maximise a programme's chances of success, efforts must be made to reduce complexity in as many NASSS domains as possible. That said, the temptation to address an oversimplified, abstracted version of the problem (in any domain) should be resisted. Bounded rationality (delineating the problem as a simple set of algorithmic decisions and defining various complicating factors as out of scope, for example) is sometimes a necessary tactic for policymakers – but it is unlikely to work in practice.

Rather than oversimplifying, we suggest that the approach to the problem should incorporate acknowledging and exploring complexity in all its richness across the multiple domains of the NASSS framework – including the condition or illness, the technology, the value proposition, the intended adopters, the organisation(s), the wider context and likely evolution of the technology and the programme-in-context over time. Next, seek to identify any sub-domains in which this complexity might be reduced. This is likely to mean scaling back on the kinds of illness or condition for which the technology is claimed to be useful; reducing the technology's interconnections (and other complex features); sharpening the value proposition; reducing the demands made on staff and patients, and proactively addressing national regulatory and policy barriers. In each of these areas for potential complexity reduction, specific theories (some of which are described above) may be relevant.

Because complexity tends to be inherent in healthcare programmes, the key challenge is often to develop ways of 'running with' complexity rather than seeking to eliminate it. The literature on complex systems suggests a number of strategies for running with complexity, including: strengthen programme leadership (and consider how to draw on distributed leadership to complement overall programme leadership); co-develop and sustain a clear and compelling vision for the programme while at the same time tolerating multi-stakeholder perspectives; identify and talk about uncertainty especially when it cannot be resolved; develop individuals and support the adaptive actions they take when implementing the programme at the front line; create incentives for delivering on broad objectives (but leave the detail to front-line staff) and provide them with slack resources (e.g. an accessible draw-down budget to use as appropriate); build relationships and manage stakeholder conflict; control programme growth (e.g. minimise scope creep); co-design pathways and work routines with intended end-users; acknowledge and respond to emergence, appreciating that unintended consequences will occur; and seek to better understand and work with the policy or regulatory context.

In conclusion, we live in a world that is saturated with technology, yet the pervasive problems of non-adoption, abandonment and failure of scale-up, spread and sustainability of technology programmes show no signs of abating. Time after time, the strategic focus is drawn narrowly to the technology and actors are seduced by over-enthusiastic sales pitch and distracted by simplistic models and metaphors (e.g. 'tipping point'). The dynamic socio-technical system into which new technologies and care practices must become embedded is overlooked or ignored – yet understanding and navigating its multiple interacting domains are key to programme success.

Teaching questions for reflection

1. How would you define complexity?
2. What are the features of a complex adaptive system?
3. Using your own example of a health informatics project, identify key areas of complexity in the following domains: the condition or illness, the technology, the value proposition, the intended adopters, the organisation(s), the wider context and the embedding and adaptation of the technology and the programme-in-context over time.
4. Using your own example, consider how these areas of complexity could be either reduced or managed.

References

[1] T. Greenhalgh, J. Wherton, C. Papoutsi, et al., Beyond adoption: A new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies, *Journal of Medical Internet Research* **19**(11) (2017), e367.

[2] T. Greenhalgh, J. Wherton, C. Papoutsi, et al., Analysing the role of complexity in explaining the fortunes of technology programmes: empirical application of the NASSS framework, *BMC Medicine* **16**(1) (2018), 66.

[3] S. Cohn, M. Clinch, C. Bunn, and P. Stronge, Entangled complexity: why complex interventions are just not complicated enough, *Journal of Health Services Research & Policy* **18**(1) (2013), 40-43.

[4] J. Braithwaite, R.L. Wears, and E. Hollnagel, *Resilient Health Care Volume 3: Reconciling Work-As-Imagined and Work-As-Done,* CRC Press, Boca Raton: FL, 2017.

[5] P.E. Plsek and T. Greenhalgh, Complexity science: The challenge of complexity in health care, *BMJ* **323**(7313) (2001), 625-8.
[6] A. Shiell, P. Hawe, and L. Gold, Complex interventions or complex systems? Implications for health economic evaluation, *BMJ: British Medical Journal* **336**(7656) (2008), 1281.
[7] P. Hawe, A. Shiell, and T. Riley, Theorising interventions as events in systems, *American Journal of Community Psychology* **43**(3-4) (2009), 267-276.
[8] T. Greenhalgh and J. Russell, Why do evaluations of eHealth programs fail? An alternative set of guiding principles, *PLoS Medicine* **7**(11) (2010), e1000360.
[9] J.M. Corbin and A. Strauss, *Unending work and care: Managing chronic illness at home,* Jossey-Bass, New York, 1988.
[10] T. Greenhalgh, R. Procter, J. Wherton, et al., What is quality in assisted living technology? The ARCHIE framework for effective telehealth and telecare services, *BMC Medicine* **13** (2015), 91.
[11] J.T. Hart, The inverse care law*, The Lancet* **297**(7696) (1971), 405-412.
[12] A.K. Schaink, K. Kuluski, R.F. Lyons, et al., A scoping review and thematic classification of patient complexity: offering a unifying framework*, Journal of Comorbidity* **2**(1) (2012), 1-9.
[13] J. Pols, *Care at a distance: On the closeness of technology,* Amsterdam University Press, Amsterdam, 2012.
[14] R.H. Coase, The nature of the firm*, Economica* **4**(16) (1937), 386-405.
[15] R.H. Coase, The problem of social cost*, The Journal of Law and Economics* **56**(4) (2013), 837-877.
[16] B.R. Theodore, J. Whittington, C. Towle, et al., Transaction cost analysis of in-clinic versus telehealth consultations for chronic pain: Preliminary evidence for rapid and affordable access to interdisciplinary collaborative consultation, *Pain Medicine* **16**(6) (2015), 1045-1056.
[17] S. Abimbola, K.N. Ukwaja, C.C. Onyedum, et al., Transaction costs of access to health care: Implications of the care-seeking pathways of tuberculosis patients for health system governance in Nigeria*, Global Public Health* **10**(9) (2015), 1060-1077.
[18] J.J. Wallis and D. North, Measuring the transaction sector in the American economy*,* 1870-1970, in *Long-term factors in American economic growth*. University of Chicago Press, Chicago (1986), 95-162.
[19] M.C. Munger, *Tomorrow 3.0: Transaction Costs and the Sharing Economy,* Cambridge University Press, Cambridge, (2018).
[20] P. Lehoux, F.A. Miller, G. Daudelin, and J.-L. Denis, Providing value to new health technology: The early contribution of entrepreneurs, investors, and regulatory agencies, *International Journal of Health Policy and Management* **6**(x) (2017), 1-10.
[21] M. van Limburg, J.E. van Gemert-Pijnen, N. Nijland, et al., Why business modeling is crucial in the development of eHealth technologies, *Journal of Medical Internet Research* **13**(4) (2011), e124.
[22] A. Bandura, *Social learning theory*, Prentice-Hall, Englewood Cliffs, 1976.
[23] T. Greenhalgh, R. Stones, and D. Swinglehurst, Choose and Book: a sociological analysis of 'resistance' to an expert system, *Social Science and Medicine* **104 (**2014), 210-219.
[24] C.R. May, D.T. Eton, K. Boehmer, et al., Rethinking the patient: using Burden of Treatment Theory to understand the changing dynamics of illness*, BMC Health Services Research* **14**(1) (2014), 1.
[25] T. Greenhalgh, G. Robert, F. Macfarlane, et al., Diffusion of innovations in service organizations: systematic review and recommendations*, Milbank Q* **82**(4) (2004), 581-629.
[26] L. Fitzgerald and A. McDermott, *Challenging Perspectives on Organizational Change in Health Care,* Routledge, London, 2016.
[27] C. May and T. Finch, Implementing, embedding, and integrating practices: an outline of normalization process theory*, Sociology* **43**(3) (2009), 535-554.
[28] E.M. Rogers, *Diffusion of innovations*, Simon and Schuster, New York, 2010.
[29] E. Hollnagel, J. Braithwaite, and R.L. Wears, *Resilient Health Care*. Ashgate Publishing, Ltd, 2013.
[30] C. Nemeth, R. Wears, D. Woods, et al., *Minding the gaps: creating resilience in health care*, in *Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 3: Performance and Tools),* K. Henriksen, J.B. Battles, and M.A. Keyes, Editors, Agency for Healthcare Quality (US): Rockville, Maryland, 2008.
[31] T. Greenhalgh, S. Shaw, J. Wherton, et al., Video outpatient consultations: A case study of real-world implementation at macro, meso, and micro level*, Journal of Medical Internet Research* (2018).
[32] A. Triantafyllidis, C. Velardo, T. Chantler, et al., A personalised mobile-based home monitoring system for heart failure: The SUPPORT-HF Study*, International Journal of Medical Informatics* **84**(10) (2015), 743-53.
[33] T. Greenhalgh and S. Shaw, *Understanding heart failure; explaining telehealth–a hermeneutic systematic review, BMC Cardiovascular Disorders* **17**(1) (2017), 156.

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

doi:10.3233/SHTI190124

Reflecting and Looking to the Future: What Is the Research Agenda for Theory in Health Informatics?

Philip J. SCOTT[a], Nicolette F. DE KEIZER[b] and Andrew GEORGIOU[c]
[a] *Centre for Healthcare Modelling & Informatics, University of Portsmouth, UK*
[a] *Department of Medical Informatics, Amsterdam UMC, University of Amsterdam, Amsterdam Public Health Institute, The Netherlands*
[c] *Centre for Health Systems and Safety Research, Macquarie University, Sydney, New South Wales, Australia*

Abstract. In this chapter, we reflect on the aim and objectives of the textbook and address known gaps in our theory coverage. We reinforce the importance of theory in health informatics and review the varying disciplinary origins of the theories considered in the book. We discuss the question of what makes a good theory and how to know which one is relevant for a given study. We recognize the limitations of the body of theory that we have presented and suggest what might be regarded as "native" theory that is original to health informatics. Finally, we propose topics to form a research agenda for theory in health informatics.

Keywords. Health informatics, Theory, Epistemology, Research

1. Introduction

We wanted this book to provide a scientific knowledge base to progress the agenda of evidence-based health informatics [4] by emphasising theory-informed work which sets out to 'enrich our understanding of this complex field' [5]. The first hurdle we confronted was that the definition of 'theory' within our field was ambiguous and needed to be broadened and made flexible so as to be "abstract enough to permit generalization, but concrete enough to permit testing". As a consequence, when predictive theories were not available, frameworks including (also non-causal) associations between concepts were selected. We have included fifteen different theories and frameworks in the book: five from information science and technology, nine from the social and psychological sciences and one ambitious framework that aims to integrate several theoretical approaches to the adoption and sustainability of health informatics interventions. Within each chapter use cases have been described that showed how the particular theory or framework enriched the understanding of the underlying mechanisms of health IT interventions to ultimately improve care. Although we do not pretend to have comprehensively covered the whole health informatics field with our fifteen chapters of theories and frameworks (see section 2), we believe that the theories included provide a solid base which will inspire further developments as needed.

We believe that enhanced understanding of applied theories in health informatics can make a positive addition to health IT research, implementation and education.

Researchers often have a role in designing health IT and evaluating its effect. A good example is the chapter by Gude and Peek which outlines how Control Theory was used to design electronic audit and feedback interventions and to understand the mechanisms behind it through thorough theory-based evaluation. When theories and frameworks appeared to be predictive in the success (or failure) of implementation of a health IT intervention, such as shown in the chapters on technology adaption of Ammenwerth and Greenhalgh et al, it seems unethical and inefficient when implementers of health IT intervention do not use this knowledge. We also believe this book provide important educational material to teachers and students. First it gives an overview of relevant theories and frameworks in the field. Second, by the example use cases young scientists may translate the use of these theories and frameworks to other applications they will encounter in their career. Third, the teaching questions at the end of each chapter support further discussion among students and teachers to deepen the understanding of theories and their applications.

We had three specific objectives with this book, discussed in the following sections of this concluding chapter:

- To show where and how interdisciplinary theories have been applied in health informatics
- To identify theory developed specifically within health informatics
- To highlight where further work is necessary to develop theory-based approaches.

The use cases in chapter two to sixteen show a wide range of applications of interdisciplinary theories in health informatics. We summarize and reflect upon this in sections 4-6 of this chapter. Researchers and implementers are motivated to add to this set of applications, thereby broadening the knowledge on applicability of theories in a variety of contexts. Researchers are encouraged to publish results, either positive or negative, so that all can learn from these findings. In section 7 we will discuss the limited amount of theories specifically developed within health informatics illustrating the fledgling status of health informatics as a discipline. In this final chapter we also offer our own overview of theory within the overall health informatics body of knowledge and propose a research agenda to contribute to the development of the health informatics discipline.

2. How comprehensive is our theory coverage?

We organised the chapters of this book according to two of the AMIA "foundational domains" of health informatics [60]: information science and social science (see Table 1). The third foundational domain, health science, has not directly featured in this book. Yet, of course, theory abounds in the health sciences. There are theories of ethics [41] and applied theory drawn from the natural sciences [34]. We think of the fundamental theories of Western medicine such as Harvey's theory of blood circulation [13] and the germ theory of disease [1], plus more recent developments such as the inflammation theory of disease [31] and social determinants theory [61]. The nursing profession has a rich and extensive body of theory, from Florence Nightingale onwards [44]. So far, in itself, the body of health sciences theory may seem less relevant for health informatics – though we return to this in section 6. We observe that many of the theories covered in this book might be positioned at the intersection of health science with the other domains. Moreover, with theory that is interdisciplinary there are inevitably different ways that ideas can be categorized. Life does not exist in neat boxes, so it is no surprise that health informatics theory and practice has a certain "messiness" [6].

Table 1. Fifteen interdisciplinary theories grouped by AMIA foundational domain [60]

AMIA Foundational Domain	Theories
Information Science and Technology	General System Theory and Process Mining Shannon's Information Theory Information Value Chain Theory User-Centred Design and Activity Theory Technology Adoption Models NASSS
Social and Behavioural Science	Distributed Cognition Actor-Network Theory Collective Mindfulness Boosting Framework Deterioration Communication Management Theory Resilient Health Care Health Behaviour Theory[1] Control Theory Normalisation Process Theory

We know that there are important topics that we have not been able to include in this volume, some simply because we could not find authors able to write within the given deadline. For example, in the Information Science and Technology section we have not covered theory related to biomedical ontologies [48], clinical information modelling [33] and healthcare information governance [38]. These would all be valuable to add in a future edition. There may also be relevant theories from computer science and statistics, especially relating to machine learning and data science more generally, that are potential candidates for inclusion. The section on Social and Behavioural Science would benefit from extensions on among others: shared mental models [57], decision theory [23], process and knowledge theory [24], practice theory [22] and team chemistry [62]. We hope this book will have regular revisions in the coming years thereby evolving the theory-base.

We invite readers to inform us about important missing applied theories in health informatics that we might include in a revised edition. Together we are responsible to mature our discipline and we believe that a theory-base is essential in this transformation.

3. Why it is important to consider theory?

"Reliance on theory demonstrates a level of sophistication in any discipline…" [11].

As in many other health care disciplines, informal theory plays an implicit role in how health informatics practitioners undertake their work. According to Davidoff et al.,

[1] This is a group of theories, not a single theory.

in their paper on "demystifying theory" the challenge is not whether theory is used, but whether the use of theory is explicated [21].

The use of theory in health informatics can help to understand phenomena, guide our analyses and improve our appreciation of the significance of research findings. Traditionally, theory has not played a dominant role in health informatics. Many researchers have noted the under-utilisation of theory to explain changes or help to predict outcomes [16]. This paucity of theory usually coincides with the absence of information about why an IT system may be useful in one context but not another [32].

Nevertheless, as this text book attests, there has been a significant uptake and growth of theory-driven approaches in health informatics particularly as a means of helping to identify concepts that are critical to understanding complex situations [11]. These developments reflect a strong push by health care administrators, software vendors, consumers, clinicians and academics to enhance understanding the outcomes of health IT systems, as a means of improving their design, implementation and sustainability into the future.

How is theory used in health informatics? It is now widely recognised that the implementation of health information systems can have a major impact on the delivery of health care and the outcomes of that care [2]. As a consequence, there is considerable attention provided to the imperative to ensure that health IT is rigorously evaluated. The utilisation of theory in health informatics has paralleled many of the developments in evaluation research over the last few decades. This is because theory can provide a frame of reference that can help us to understand the significance of evaluation findings [15]. In this way health informatics is able to go beyond a simple "black box" evaluation which may tell us whether or not a health IT system works, towards a greater appreciation of the underlying causal mechanisms and context in which it is placed.

This textbook did not intend to identify all the relevant theories currently used in health informatics. To our knowledge, there are not many comprehensive examinations of the breadth of the use of theory in health informatics [51]. This textbook provides a wide sample of theories currently employed within the health informatics discipline. Our choice of theories was purposive and informed by an intensive and iterative engagement with our health informatics colleagues over many years, including through discussions at workshops and panels at major health informatics events. Health informatics research draws from many other theories in other fields and disciplines including, sociology, psychology, information systems, implementation science and communication. Table 2 identifies some of the key disciplinary origins (drawing on Greenhalgh et al's meta-narrative on electronic patient record research [28]) of the theories presented in this textbook along with their scope and existing areas of utilisation.

Table 2. The origins, scope and utilisation of the interdisciplinary theories in this textbook.

Theory	**Utilisation**	**Disciplinary origins**	**Scope**
General System Theory and Process Mining	Learning health systems; Process mining of care pathways and simulation.	Complexity theory; computer science	The use of big data analytics to develop better, integrated and personalised pathways of care for patients.
Shannon's Information Theory	Assisting choice of diagnostic tests in a clinical setting; identification of redundancies in clinical tests.	Computer science; software engineering; molecular biology; statistical inference; natural language processing.	Informing medical decision making.
Information Value Chain Theory	System design; national summary EHRs; audit and feedback;	Business studies; psychology; computer science; management.	How organisations adopt and assimilate information systems.
User Centred Design and Activity Theory	Design of a mobile health IT system to improve healthcare delivery; Evaluation of a mHealth system used by community health workers.	Sociology; psychology; ergonomics; computer science; anthropology; Software engineering	Complexities of users and their interaction with computer systems.
Technology Acceptance Models	Usefulness and ease of use of EHRs among nurses; home telehealth acceptance among older people.	Evidence-based medicine; Computer science.	The benefits of digital health and how to achieve them.
Distributed Cognition	Situational awareness in cardiac surgery; Handovers in psychiatric emergency; infection control information; safety; pharmacy.	Organisational sociology; social psychology, philosophy	How social structures recursively shape and are shaped by human agency and the role of technology.
Actor-Network Theory	Medical records; IT system failures; NPfIT implementation;	Philosophy, sociology, linguistics.	Study of socio-technical networks and what emerges from these.
Collective Mindfulness	Project risk management; Adaptation of digital health systems.	Management, sociology, social psychology, anthropology.	How organisational members make sense of information systems and assimilate them.
Boosting Framework	Design of patient decision aids. Communication of risk.	Sociology; philosophy; social psychology; management.	Application of boosting theory to foster choices and shared decision making.
Deterioration Communication Management Theory	Information transfer; ICT evaluation; Junior doctor training; Care for deteriorationg patients.	Management, sociology, social psychology, anthropology.	How organisational members make sense of information systems and assimilate them.

Theory	**Utilisation**	**Disciplinary origins**	**Scope**
Resilient Health Care	Video consultation and triage service; Hospital response to an unexpected event.	Systems and management research; resilience engineering.	The way that systems cope successfully with unwanted outcomes (or events) that are unexpected.
Health Behaviour Theory	Smoking cessation; Choice of breast cancer therapy; Exercise.	Evidence-based medicine; social psychology; management	How to achieve organisational level change in health care.
Control Theory	Audit and feedback interventions; diabetes management; behaviour change techniques; blood transfusion practice; pain management in intensive care.	Computer science; psychology; business studies.	Self-regulation and human behaviour; why interventions were or were not successful.
Normalization Process Theory	Evaluation of a digital health intervention for Type 2 diabetes; Preoperative information system within a surgical pre-assessment clinic.	Social psychology; management	Why are new technologies and working practices implemented successfully in some settings but not in others?
NASSS Framework	Telehealth system for heart failure.	Evidence-based medicine; sociology; management; social psychology; systems and management research	The multiple influences on a complex project; how complexity might be reduced and how individuals and organisations might be supported to handle complexities.

4. What makes up a good theory and which theory should be used?

Different disciplines tend to have their own perspective of what constitutes a theory and the criteria for a good-quality theory [21]. Whilst there are some criteria that are likely to be universal (e.g., clarity of concepts, causality, testability, generalisability etc.), there are also likely to be other criteria specific to the needs of health informatics. We expect that a better appreciation of what exists in terms of theory will help to spark more attention to what constitutes a good health informatics theory. Evidence from related disciplines (e.g., implementation science, quality improvement) who face some parallel challenges to health informatics, suggest that the choice of theory is often made on an arbitrary basis, and usually based on expediency or previous exposure [7]. Clearly, this question merits attention in a research agenda for theory in health informatics.

There are different approaches to how theory is developed and tested. For instance "adaptive" theory can be defined as a combination of pre-existing theory and incoming evidence [35; 36]. Adaptive theory approaches can thus be shaped by research evidence, even while the pre-existing theoretical material (framework, concept) is helping to shape the course of evidence gathering. Alternatively, "grounded theory" is based on the notion that theory emerges from the research data. For grounded theory [26; 36] one of the measures for judging the relevance of a theory is whether or not it is comprehensible to the subjects of the research.

5. Limitations

In addition to the limitations of scope discussed in section 2, it is helpful also to reflect on the explicit limitations of each theory. Table 3 summarises the limitations identified in each chapter by the authors.

Table 3. Explicit limitations of the interdisciplinary theories in this textbook.

Theory	Limitations
General System Theory and Process Mining	"Anything could be seen as a system depending on the boundaries you set".
Shannon's Information Theory	Need to explicitly model the "noise" that is inherent in the communication model. Shannon entropy, relative entropy and conditional entropy are non-intuitive concepts.
Information Value Chain Theory	Relatively new, with few applications. The theory does not attempt to provide detailed mechanistic explanations for the impact of information technology beyond the causality implied in the structure of the chain itself. As with any theory that relies on quantitative measurements, it is important to ensure that data used in any analysis actually measures what it is meant to.
User Centred Design and Activity Theory	Software application needs to collect and infer relevant contexts to understand the user's situation. Users will invariably have different perceptions, understanding and expectations, influenced by social, cultural and historical context.

Theory	Limitations
Technology Acceptance Models	A number of TAM extensions have been proposed to overcome some limitations in the original model.
Distributed Cognition (DiCoT)	Distributed cognition encourages a level of description about a system or process that lends itself to developing design ideas, but it may not readily emphasise the role of individuals or emotions as it focuses on systems and more observable functional issues.
Actor-Network theory	Lack of predictive power. Not internally consistent. Lacks specificity. Treatment of human actors and non-human actors as equal. Terminology is only loosely defined.
Collective Mindfulness	Rarely applied in health informatics. Recommendations may be difficult to put into practice. Principles could be viewed as ideals rather than descriptors.
Boosting Framework	Rarely applied in health informatics. It is a framework not theory and it helps to explicate some guiding principles for future research, from which testable assumptions can be derived. The boosting framework does not yet provide a full-blown process model with detailed "how-to" information describing how research evidence can be translated into practical health informatics solutions.
Deterioration Communication Management Theory	Classical Grounded Theory (CGT) focuses on one main concern, unlike constructivist Grounded Theory (GT) which aims to understand multiple perspectives in a social process. CGT has seldom been used in information systems research.
Resilient Health Care (RHC)	Resilient Health Care theory is relatively new and many of its tools are still in their infancy. RHC tools may be used to complement determinant frameworks such as computational simulation modelling.
Health Behaviour Theory	The use of behaviour change theory in health informatics interventions originates mainly from other disciplines: psychologists and public health workers familiar with behaviour change theory. Many interventions developed by people working in health informatics do not report using a health behaviour or behaviour change theory.
Control Theory (CT)	In the Audit & Feedback literature it has often been used but not explicitly reported. HI interventions are typically complex and placed into a social and organisational context. This context is not in the scope of CT. It provides no guidance as to which factors related to the context, recipients, or feedback itself may influence success of the feedback loop.
NASSS Framework	Published studies about the application of the NASSS framework are limited. Key challenge is to find ways of "running with" complexity, instead of seeking to "eliminate it".

6. Does health informatics have any theory of its own?

So far, we have focused on the first of our three objectives: to show where and how interdisciplinary theories have been *applied* in health informatics. Most of the theories have been *developed* in other fields. We now turn to the question of theory developed specifically within health informatics, to consider where further work is necessary to develop theory-based approaches.

We suggest that there are only three examples of "native" health informatics theory in the textbook:

- Distributed Cognition (as it is specifically about "information processing in sociotechnical systems")
- Deterioration Communication Management Theory (as its aim is "to improve the design and implementation of ICT systems for communication to and from junior hospital doctors")
- The NASSS Framework (as it is focused explicitly on "technologies in health and care organizations").

What else is out there?

Arguably, the oldest theories in health informatics are the "determinant frameworks" (in Nilsen's terminology [45]) relating to the structure and content of patient records. This is unsurprising as it is perhaps the most obvious overlap between healthcare and information. In 1605, Francis Bacon harked back to the narrative case histories of the school of Hippocrates as the ideal [20]. Later, Thomas Sydenham, the 'English Hippocrates', wrote in 1676 that an "exact history" of every case of disease would improve therapy by making it empirically obvious how to proceed [20]. Francis Clifton proposed to the Royal Society in 1731 that medical observations should be recorded in a particular tabular format to simplify record-keeping and facilitate comparative analysis [50]. In the 1960s, Larry Weed famously proposed problem-oriented medical records "to guide and teach" [63; 64] and this approach has been adopted in some electronic health record systems. Recent health informatics work has included the development of detailed clinical information models of re-usable concepts in representations such as archetypes [43] and Fast Healthcare Interoperability Resources (FHIR) [58]. While that modelling work is for the purpose of technical implementation not informatics theory, there is still an implied hypothesis that such shared concepts are sufficiently stable, definable and comprehensible to be safe and meaningful as a common language of healthcare.

The most basic theory in modern health informatics seems to be Friedman's fundamental theorem [25]. Friedman asserted that "A person working in partnership with an information resource is 'better' than that same person unassisted", with three important corollaries: (1) That informatics is more about people than technology; (2) In order for the theorem to hold, the resource must offer something that the person does not already know; and (3) Whether the theorem holds depends on an interaction between person and resource, the results of which cannot be predicted in advance. The theorem has been questioned [40] and modifications to the wording have been suggested [30; 39], but the common sense of Friedman's theorem seems generally accepted.

Another quite basic proposition is the "first law" that van der Lei proposed: "data shall be used only for the purpose for which they were collected". The continuing validity of this has been questioned [54] and it is expressed as a normative principle rather than an explanatory model or predictive hypothesis, but there is an implicit prediction that if the “law” is not followed then the conclusions from the data will be flawed.

We have not conducted a systematic review, but in the preparation of this textbook we have informally reviewed a broad range of literature and in Table 4 we offer an illustrative sample of contributions that might be regarded as theory in health informatics.

Table 4. A sample of candidate theories in health informatics.

Reference	What is the “theory”?	What does it claim to explain?
[3]	Thematic Hierarchical Network Model for Computerised Physician Order Entry (CPOE) Consequences.	Relationship between categories of unintended consequences of CPOE.
[4]	Sociological perspective on EHR design.	That medical work is not a linear rational process, so EHR design should rather support fluidity of knowledge and collaborative, interactive working.
[5]	Human Factors Engineering (HFE) approach to biomedical informatics applications for healthcare.	That HFE shows why implementations are successful or not.
[8; 9]	Cognitive span of the process of clinical diagnosis.	That it is at the latter end of the diagnostic process that decisions become algorithmic and therefore when computers become potentially useful.
[12]	Clinical domain reference ontologies.	The ideal features and attributes of reference ontologies for a specific clinical knowledge domain.
[17]	Thematic synthesis of controlled medical vocabulary requirements.	The ideal features and attributes of a computable controlled medical vocabulary.
[18]	Alternative paradigm for modelling clinical interactions based on psychological concept of “common ground”.	That the typical computational model of communication does not correspond with actual clinical experience of mostly interrupt-driven human interaction.
[46]	Three general principles to determine whether CPOE implementation will succeed.	Why CPOE implementations succeed or not.
[47]	A nine-factor construct of clinician perceptions about computerized protocols.	How clinicians react to computerized protocols.
[49]	Evaluation model of clinical information systems viewed from health system perspective rather than functional or organizational assessment.	That the full picture of time effects of clinical systems can only be evaluated at whole-system level not just by unit component effects.
[55; 56]	Architecture for sharing EHRs independently of disparate healthcare providers.	That independent health record banks offer a more sustainable solution for lifetime EHRs than records held by providers, payers or government agencies.

As noted in section 3, the development or application of theory in health informatics is predominantly implicit rather than explicit. Readers (and editors of theory textbooks!) are usually left to infer the theoretical contribution. The STARE-HI guideline for reporting of evaluation studies in health informatics [10] includes "theoretical background" as a section in Methods – this is very far from routinely followed. We certainly would not want to see our field adopt an extreme position where theory is idolized and academic papers become weighed down with ponderous and pretentious intellectual displays, as has been reported in the field of management [29], but health informatics generally appears to be at the opposite end of that spectrum and needs a nudge towards a stronger theoretical approach.

7. A research agenda for theoretical health informatics

Finally, we consider theoretical topics for future research that have been identified in the literature and some that we propose based on our learning in preparing this textbook.

From Table 4, we offer some specific areas for consideration:

- Theory of CPOE implementation
- Theory of sociological design of EHRs
- Theory of computational diagnostic support
- Theory of clinical communication patterns
- Theory of healthcare protocol adoption
- Theory of systemic evaluation
- Theory of personally controlled electronic health records.

The converging paradigms of precision medicine, Learning Health Systems and implementation science seem to offer a particularly fruitful ground for theory development given the central role of informatics in each of these fields [14; 52; 59]. A key aspect of this convergence is clinical decision support, which has long been an important area of study in health informatics [27; 42], though the robustness of its evaluation still has a way to go in terms of scientific measurement practice [53]. The need for sound theoretical foundations for this work has been recognised [23; 24] but in some quarters seems to be perceived as merely a technical implementation challenge. We argue that this is a prime area where we should expect to see emerging theory.

In addition to these specific topics, there is the general lack of replication studies in health informatics [19]. Without such a culture of replication studies, our field will be dominated by single-case evaluations that do not lend themselves to broader theoretical generalisation. Theories may not lend themselves to the same form of replication of findings but there is still a need for validation. Theoretical approaches often take a triangulation approach or utilise member validation methods.

We have already noted in section 4 that criteria need to be developed for selecting relevant theory and assessing theory quality. As suggested in [7], there should be transparent reporting of the criteria used to select theories in research studies. This implies the need for a comprehensive list of criteria that are used to choose a theory. Such an approach in health informatics would encourage reflective thinking, explication and generalisability. The theory criteria would include the identification of key constructs; informing data collection; enhancing conceptual clarity; clarifying terminology and hypothesising relationships [7]. This would be a useful contribution to the field and could inform the revision of existing reporting guidelines [10].

8. Conclusion

We thank our readers for joining us on this journey into the theoretical side of health informatics. We have learned much and enjoyed developing the book. We hope you find the book useful and welcome suggestions for a more extensive second edition. Finally, we hope that you are now part of the health informatics community that agrees with Lewin that: "there is nothing more practical than a good theory" [37].

Teaching questions for reflection

1. In your area of health informatics practice, what would you identify as the most useful interdisciplinary theories in this textbook?
2. What additional theories can you propose for a future edition of the textbook?
3. How would you evaluate the scientific maturity of health informatics, based on its current approach to theory?
4. What do you feel are the priorities for theoretical topics in health informatics that need further research?
5. What would criteria for health informatics theory selection look like?

References

[1] The history of the germ theory, *British Medical Journal* **1** (1888), 312-313.
[2] E. Ammenwerth, J. Brender, P. Nykanen, H.U. Prokosch, M. Rigby, and J. Talmon, Visions and strategies to improve evaluation of health information systems. Reflections and lessons based on the HIS-EVAL workshop in Innsbruck, *Int J Med Inform* **73** (2004), 479-491.
[3] J.S. Ash, D.F. Sittig, R.H. Dykstra, K. Guappone, J.D. Carpenter, and V. Seshadri, Categorizing the unintended sociotechnical consequences of computerized provider order entry, *Int J Med Inform* **76 Suppl 1** (2007), S21-27.
[4] M. Berg, Medical work and the computer-based patient record: a sociological perspective, *Methods Inf Med* **37** (1998), 294-301.
[5] M.C. Beuscart-Zephir, P. Elkin, S. Pelayo, and R. Beuscart, The human factors engineering approach to biomedical informatics projects: state of the art, results, benefits and challenges, *Yearb Med Inform* (2007), 109-127.
[6] E. Billo and N. Hiemstra, Mediating messiness: expanding ideas of flexibility, reflexivity, and embodiment in fieldwork, *Gender, Place & Culture* **20** (2013), 313-328.
[7] S.A. Birken, B.J. Powell, C.M. Shea, E.R. Haines, M. Alexis Kirk, J. Leeman, C. Rohweder, L. Damschroder, and J. Presseau, Criteria for selecting implementation science theories and frameworks: results from an international survey, *Implementation Science* **12** (2017), 124.
[8] M. Blois, *Information and Medicine: Nature of Medical Descriptions*, University of California Press, Berkeley CA, 1984.
[9] M.S. Blois, Clinical judgment and computers, *N Engl J Med* **303** (1980), 192-197.
[10] J. Brender, J. Talmon, N. de Keizer, P. Nykanen, M. Rigby, and E. Ammenwerth, STARE-HI - Statement on Reporting of Evaluation Studies in Health Informatics: explanation and elaboration, *Appl Clin Inform* **4** (2013), 331-358.
[11] P.F. Brennan, Standing in the shadows of theory, *Journal of the American Medical Informatics Association* **15** (2008), 263-264.
[12] A. Burgun, Desiderata for domain reference ontologies in biomedicine, *J Biomed Inform* **39** (2006), 307-313.
[13] J.J. Bylebyl, The growth of Harvey's "de motu cordis", *Bulletin of the History of Medicine* **47** (1973), 427-470.
[14] D.A. Chambers, W.G. Feero, and M.J. Khoury, Convergence of Implementation Science, Precision Medicine, and the Learning Health Care System: A New Model for Biomedical Research, *JAMA* **315** (2016), 1941-1942.

[15] H.T. Chen, *Theory-driven evaluations*, Sage, London, 1990.
[16] M. Chiasson, M. Reddy, B. Kaplan, and E. Davidson, Expanding multi-disciplinary approaches to healthcare information technologies: What does information systems offer medical informatics? , *International Journal of Medical Informatics* **76** (2007), S89-S97.
[17] J.J. Cimino, Desiderata for controlled medical vocabularies in the twenty-first century, *Methods Inf Med* **37** (1998), 394-403.
[18] E. Coiera, When conversation is better than computation, *J Am Med Inform Assoc* **7** (2000), 277-286.
[19] E. Coiera, E. Ammenwerth, A. Georgiou, and F. Magrabi, Does health informatics have a replication crisis?, *J Am Med Inform Assoc* **25** (2018), 963-968.
[20] A. Cunningham, The transformation of Hippocrates in seventeenth century Britain, in: *Reinventing Hippocrates,* D. Cantor, ed., Ashgate, Aldershot, 2002.
[21] F. Davidoff, M. Dixon-Woods, L. Leviton, and S. Michie, Demystifying theory and its use in improvement, *BMJ Qual Saf* **24** (2015), 228-238.
[22] M.S. Feldman and W.J. Orlikowski, Theorizing practice and practicing theory, *Organization Science* **22** (2011), 1240-1253.
[23] J. Fox, Cognitive systems at the point of care: The CREDO program, *J Biomed Inform* **68** (2017), 83-95.
[24] J. Fox, D. Glasspool, V. Patkar, M. Austin, L. Black, M. South, D. Robertson, and C. Vincent, Delivering clinical decision support services: there is nothing as practical as a good theory, *J Biomed Inform* **43** (2010), 831-843.
[25] C.P. Friedman, A "fundamental theorem" of biomedical informatics, *J Am Med Inform Assoc* **16** (2009), 169-170.
[26] B.G. Glaser and A.L. Strauss, *The discovery of grounded theory: Strategies for qualitative research*, Aldine de Gruyter, New York, 1967.
[27] R.A. Greenes, D.W. Bates, K. Kawamoto, B. Middleton, J. Osheroff, and Y. Shahar, Clinical decision support models and frameworks: Seeking to address research issues underlying implementation successes and failures, *J Biomed Inform* **78** (2018), 134-143.
[28] T. Greenhalgh, H.W. Potts, G. Wong, P. Bark, and D. Swinglehurst, Tensions and paradoxes in electronic patient record research: a systematic literature review using the meta-narrative method, *Milbank Q* **87** (2009), 729-788.
[29] D.C. Hambrick, The field of management's devotion to theory: Too much of a good thing?, *Academy of Management Journal* **50** (2007), 1346-1352.
[30] J.S. Hunter, Enhancing Friedman's "fundamental theorem of biomedical informatics", *J Am Med Inform Assoc* **17** (2010), 112; author reply 112-113.
[31] P. Hunter, The inflammation theory of disease: The growing realization that chronic inflammation is crucial in many diseases opens new avenues for treatment, *EMBO reports* **13** (2012), 968-970.
[32] B. Kaplan, Evaluating informatics applications - some alternative approaches: theory, social interactionism, and call for methodological pluralism., *International Journal of Medical Informatics* **64** (2001), 39-56.
[33] S. Koch, P. Chhanabhai, P.S. Ruotsalainen, B.G. Blobel, A.V. Seppälä, H.O. Sorvari, and P.A. Nykänen, A Conceptual Framework and Principles for Trusted Pervasive Health, *J Med Internet Res* **14** (2012).
[34] R.E. Kohler Jr, The enzyme theory and the origin of biochemistry, *Isis* **64** (1973), 181-196.
[35] D. Layder, *The realist image in social science*, Macmillan, Basingstoke, 1990.
[36] D. Layder, *Sociological practice: Linking theory and social research*, Sage, 1998.
[37] K. Lewin, *Field theory in social science: selected theoretical papers*, Harper, New York, NY, 1951.
[38] S.T. Liaw, C. Pearce, H. Liyanage, G.S. Liaw, and S. de Lusignan, An integrated organisation-wide data quality management and information governance framework: theoretical underpinnings, *Inform Prim Care* **21** (2014), 199-206.
[39] S. Mani, Note on Friedman's 'fundamental theorem of biomedical informatics', *J Am Med Inform Assoc* **17** (2010), 614.
[40] V. Maojo and C.A. Kulikowski, Note on Friedman's 'what informatics is and isn't', *J Am Med Inform Assoc* **20** (2013), e365-366.
[41] E. Marseille and J.G. Kahn, Utilitarianism and the ethical foundations of cost-effectiveness analysis in resource allocation for global health, *Philos Ethics Humanit Med* **14** (2019), 5.
[42] B. Middleton, D.F. Sittig, and A. Wright, Clinical Decision Support: a 25 Year Retrospective and a 25 Year Vision, *Yearb Med Inform* **Suppl 1** (2016), S103-116.
[43] D. Moner, J.A. Maldonado, and M. Robles, Archetype modeling methodology, *J Biomed Inform* **79** (2018), 71-81.
[44] F. Nightingale, *Notes on nursing: What it is, and what it is not*, Harrison, London, 1859.
[45] P. Nilsen, Making sense of implementation theories, models and frameworks, *Implement Sci* **10** (2015), 53.
[46] A. Ozdas and R.A. Miller, Care provider order entry (CPOE): a perspective on factors leading to success or to failure, *Yearb Med Inform* (2007), 128-137.

[47] S. Phansalkar, C.R. Weir, A.H. Morris, and H.R. Warner, Clinicians' perceptions about use of computerized protocols: a multicenter study, *Int J Med Inform* **77** (2008), 184-193.
[48] D.M. Pisanelli, ed., *Ontologies in Medicine*, IOS Press, Amsterdam, 2004.
[49] L. Poissant, J. Pereira, R. Tamblyn, and Y. Kawasumi, The impact of electronic health records on time efficiency of physicians and nurses: a systematic review, *J Am Med Inform Assoc* **12** (2005), 505-516.
[50] A. Rusnock, Hippocrates, Bacon and medical meteorology at the Royal Society, 1700-1750., in: *Reinventing Hippocrates,* D. Cantor, ed., Ashgate, Aldershot, 2002.
[51] P. Scott, J. Briggs, J.C. Wyatt, and A. Georgiou, How important is theory in health informatics? A survey of UK academics, in: *User Centred Networked Health Care - Proceedings of MIE 2011,* A. Moen, S. Andersen, J. Aarts, and P. Hurlen, eds., IOS Press, Amsterdam, 2011, pp. 223-227.
[52] P. Scott, R. Dunscombe, D. Evans, M. Mukherjee, and J. Wyatt, Learning health systems need to bridge the 'two cultures' of clinical informatics and data science, *J Innov Health Inform* **25** (2018), 126-131.
[53] P.J. Scott, A. Brown, T. Adedeji, J. Wyatt, A. Georgiou, E. Eisenstein, and C. Friedman, A review of measurement practice in studies of clinical decision support systems 1998-2017, *J Amer Med Info Assoc* (In press).
[54] P.J. Scott, M. Rigby, E. Ammenwerth, J.B. McNair, A. Georgiou, H. Hypponen, N. de Keizer, F. Magrabi, P. Nykanen, W.T. Gude, and W. Hackl, Evaluation Considerations for Secondary Uses of Clinical Data: Principles for an Evidence-based Approach to Policy and Implementation of Secondary Analysis, *Yearb Med Inform* **26** (2017), 59-67.
[55] A. Shabo, A global socio-economic-medico-legal model for the sustainability of longitudinal electronic health records. Part 1, *Methods Inf Med* **45** (2006), 240-245.
[56] A. Shabo, A global socio-economic-medico-legal model for the sustainability of longitudinal electronic health records. Part 2, *Methods Inf Med* **45** (2006), 498-505.
[57] S.W. Smith and R. Koppel, Healthcare information technology's relativity problems: a typology of how patients' physical reality, clinicians' mental models, and healthcare information technology differ, *J Am Med Inform Assoc* **21** (2014), 117-131.
[58] H.R. Solbrig, E. Prud'hommeaux, G. Grieve, L. McKenzie, J.C. Mandel, D.K. Sharma, and G. Jiang, Modeling and validating HL7 FHIR profiles using semantic web Shape Expressions (ShEx), *J Biomed Inform* **67** (2017), 90-100.
[59] W.W. Stead, J.R. Searle, H.E. Fessler, J.W. Smith, and E.H. Shortliffe, Biomedical informatics: changing what physicians need to know and how they learn, *Acad Med* **86** (2011), 429-434.
[60] A.L. Valenta, E.S. Berner, S.A. Boren, G.J. Deckard, C. Eldredge, D.B. Fridsma, C. Gadd, Y. Gong, T. Johnson, J. Jones, E.L. Manos, K.T. Phillips, N.K. Roderer, D. Rosendale, A.M. Turner, G. Tusch, J.J. Williamson, and S.B. Johnson, AMIA Board White Paper: AMIA 2017 core competencies for applied health informatics education at the master's degree level, *J Am Med Inform Assoc* **25** (2018), 1657-1668.
[61] N. Valentine and O. Solar, The Conceptual Framework for Social Determinants of Health: which theory is the basis for a tool for Health Impact Assessment?, *World Health Organization* (2011).
[62] S.M.J. Vickberg and K.I.M. Christfort, Pioneers, Drivers, Integrators & Guardians, *Harvard Business Review* **95** (2017), 50-56.
[63] L.L. Weed, Medical records that guide and teach, *N Engl J Med* **278** (1968), 652-657.
[64] L.L. Weed, Medical records that guide and teach, *N Engl J Med* **278** (1968), 593-600.

Applied Interdisciplinary Theory in Health Informatics
P. Scott et al. (Eds.)

Subject Index

activity theory 49
actor-network theory 87
assessment 64
attitude to computers 64
behaviour change theory 146
boosting 109
change management 171
clinical 146
clinical audit 159
clinical deterioration 122
complex adaptive system 134
complexity of innovations 193
conditional entropy 23
continuous quality management 159
decision psychology 109
decision support systems 146
developer-user co-design 171
DiCoT 75
diffusion of innovation 193
distributed cognition 75
epistemology 205
evaluation 1, 35
feedback 159
general system theory 11
health behavior 146
health informatics 1, 64, 205
health information technology 87
health psychology 146
hospital communication systems 122
human computer interaction 49
implementation science 171
informatics 75
information technology 35
intention 64
learning health systems 11
medical diagnosis 23
mHealth 49
mindfulness 98
models 64
mutual information 23
NASSS framework 193
organisations 98
organizational behavior change 171
Patient Decision Aids (PtDAs) 109
patient safety 122, 134
personal health choices 109
process evaluation 171
process mining 11
professionalism 1
quality of health care 159
relative entropy 23
research 205
resilient health care 134
safety-I 134
safety-II 134
scale-up 193
scientific methods 1
self-regulation 159
sensemaking 98
Shannon entropy 23
Shared Decision Making (SDM) 109
smartphone 122
sociotechnical 75, 87
technology 64
technology adoption 98
text messaging 122
theoretical 64
theory 1, 205
theory-based design 146
user centred design 49
utility 35
value of information 35
value proposition 193

P. Scott et al. (Eds.)

Author Index

Abimbola, S. 193
Ammenwerth, E. 64
Blandford, A. 75
Bracher, M. 171
Braithwaite, J. 134
Clay-Williams, R. 134
Coiera, E. 35
Cresswell, K. 87
Day, K. 122
de Keizer, N.F. vii, 205
de Vries, M. 109
Franklin, B.D. 75
Friedman, C.P. v
Furniss, D. 75
Garfield, S. 75
Georgiou, A. vii, 205
Good, A. 49
Greenhalgh, T. 193
Gude, W.T. 159
Holland, R. 109
Husson, F. 75
Jansen, J. 109
Johnson, O. 11
Krause, P. 23
Liang, J. 122
Lichtner, V. 98
May, C.R. 171
Medlock, S. 146
Omisade, O. 49
Orr, M. 122
Peek, N. 159
Scott, P.J. vii, 205
van Weert, J. 109
Warren, J. 122
Westbrook, J.I. 98
Wyatt, J.C. 1, 146

The American Medical Informatics Association (AMIA) defines the term biomedical informatics (BMI) as:

> The interdisciplinary field that studies and pursues the effective uses of biomedical data, information, and knowledge for scientific inquiry, problem solving and decision making, motivated by efforts to improve human health.

This book: *Applied Interdisciplinary Theory in Health Informatics: A Knowledge Base for Practitioners*, explores the theories that have been applied in health informatics and the differences they have made. The editors, all proponents of evidence-based health informatics, came together within the European Federation of Medical Informatics (EFMI) Working Group on Health IT Evaluation and the International Medical Informatics Association (IMIA) Working Group on Technology Assessment and Quality Development. The purpose of the book, which has a foreword by Charles Friedman, is to move forward the agenda of evidence-based health informatics by emphasizing theory-informed work aimed at enriching the understanding of this uniquely complex field. The book takes the AMIA definition as particularly helpful in its articulation of the three foundational domains of health informatics: health science, information science, and social science and their various overlaps, and this model has been used to structure the content of the book around the major subject areas.

The book discusses some of the most important and commonly used theories relevant to health informatics, and constitutes a first iteration of a consolidated knowledge base that will advance the science of the field.

ISBN: 978-1-61499-990-4

ISBN 978-1-61499-990-4 (print)
ISBN 978-1-61499-991-1 (online)
ISSN 0926-9630 (print)
ISSN 1879-8365 (online)