



SciModeler: A Toolbox for Consolidating Scientific Knowledge within the Field of Health Behavior Change

Raoul Nuijten¹ · Pieter Van Gorp¹

Received: 20 September 2021 / Accepted: 10 October 2022 / Published online: 11 November 2022
© The Author(s) 2022

Abstract

Science aims to build and advance general theories from empirical data. This process is complicated by the immense volume of empirical data and scientific theories in some domains, for example in the field of health behavior change. Especially, a systematic mapping between empirical data and theoretical constructs is lacking. We propose a toolbox to establish that mapping. We adopted a modeling approach based on literature surveys to elicit requirements and to derive a metamodel. We adopted a graph-based database system to implement the metamodel, and designed a web-based tool for importing data from annotated text documents. To evaluate that toolbox (named *SciModeler*), we have conducted a case study within the field of health behavior change to record three scientific theories, three empirical studies, and the mapping in-between. We have documented how *SciModeler* aids closing gaps between empirical data and theoretical constructs. We have demonstrated that this enables new types of analyses by sharing example queries for (1) refining scientific theories, (2) exploring promising intervention strategies for a specific context, and (3) checking the potential impact of an intervention platform in a specific context. Our supplementary materials promote replication of these results. *SciModeler* can support the consolidation of scientific knowledge in the field of health behavior change, and we suggest that it may be applied within other fields, as well. An important direction for future work is promoting online collaboration on *SciModeler* graphs.

Keywords Metamodel · Theory-building · Health behavior change

Introduction

Over the past decade, research on mobile health (mHealth) apps has expanded significantly and these apps have proven to be able to promote healthier lifestyles and prevent welfare diseases [1, 2]. To promote health behaviors mHealth apps employ different intervention strategies, or behavior change techniques [3]. The effectiveness of an mHealth app to promote health behaviors largely depends on the specific (combination of) intervention strategies that the app employs [4, 5]. Oftentimes, mHealth apps are quite complex and involve many interacting intervention strategies [3–5]. However,

despite the large volume of empirical studies from the health behavior change field, it remains challenging to distinguish which intervention strategies are needed to change different behaviors, for whom, and under which conditions these are most effective [6].

One of the reasons why this challenge prevails is the ongoing scientific debate on a standardized classification of intervention strategies. In particular, mHealth research relies largely on knowledge from the field of health promotion and behavior change. That scientific domain is a blend of psychology, behavioral economics, environmental planning, urban planning, epidemiology, public policy, information technology, and computer science. With influences from so many backgrounds, there is a plethora of (overlapping) intervention strategies available [7]. Although a lot of effort has been put into deriving taxonomies of intervention strategies (e.g., see [3, 8, 9]), it has proven extremely challenging to reach consensus on a standardized taxonomy with so many scientific disciplines being involved [3, 7, 10], and this disagreement still hinders collaboration among disciplines [7]. The lack of

This article is part of the topical collection “Model-Driven Engineering and Software Development” guest edited by Slimane Hammoudi and Luis Ferreira Pires.

✉ Raoul Nuijten
r.c.y.nuijten@tue.nl

¹ Department of Industrial Engineering, Eindhoven University of Technology, Groene Loper 3, Eindhoven, The Netherlands

standardization has led to poor replicability of studies and has complicated comparisons between studies [7].

Another reason why it remains difficult to accurately estimate the impact of an intervention strategy resides in the strategy's interdependence with contextual factors. In particular, although an intervention strategy may have been demonstrated to be effective in one intervention context, these results do not automatically translate to another context [4, 5]. The current taxonomies of intervention strategies do typically not distinguish between different intervention contexts, nor do they include evidence of impact or effectiveness within a specific context [11]. Hence, while they include effective strategies for promoting user engagement, they may also contain ineffective, or even counter-effective, techniques [11]. Another approach has been to create an all-encompassing theory of behavior change (e.g., see the Integrative Model of Behavioral Prediction [12] or the COM-B System [13]). Although these theoretical frameworks help advocate general principles of health behavior change interventions, they cannot fully capture the richness of contextual factors [14]. As a result, it typically can hardly be derived for whom, and under which conditions, a specific intervention strategy is most effective [6].

This article does not aim to contribute an all-encompassing theory of health behavior change, nor a standardized taxonomy of intervention strategies. Instead, this article discusses the development and evaluation of a novel toolbox, *SciModeler*, that empowers researchers to efficiently consolidate scientific knowledge, including: (1) recording study findings and contexts in a knowledge representation that facilitates querying, (2) mapping study outcomes with theoretical constructs to refine scientific theory, and (3) making replicable predictions on the impact of a particular intervention strategy in a specific context, based on actual empirical data.

The remainder of this article is structured as follows: the next section surveys the literature on existing approaches to consolidate scientific data. The section “[Methods](#)” details our modeling approach and evaluation strategy. The section “[Results](#)” presents the corresponding results. First, it describes the modeling outcomes (i.e., literature-based requirements and metamodel). Then, it evaluates these by means of a case study that involves three scientific theories and three empirical studies from the field of health behavior change. This section comes with various supplements (ranging from an open-source annotation tool to downloadable copies of all graphs and queries). Finally, in the section “[Discussion](#)”, we discuss the principal results and derive guidelines for future work, especially also looking at limitations that still need to be overcome.

Related Work

There has been a vast amount of work related to capturing encyclopedic and factual knowledge in ontologies (e.g., by Google, Bing, IBM, BBC, or Thomson Reuters), but relatively little work focuses on representing the information contained inside scientific publications semantically [15]. Nevertheless, recent advances in Natural Language Processing (NLP) and Machine Learning (ML) have enabled the automated construction of semantic models from scientific articles [16]. Hence, these techniques could in theory be employed for the automated consolidation of knowledge of a given scientific field. However, such approaches build models that are yet unable to accurately represent argumentation and scholarly knowledge evolution in knowledge graphs [15], because argumentation requires understanding cause and effect (i.e., an operation that NLP and ML systems can hardly perform, since these systems cannot intervene in the world [17]). Hence, in this section, we discuss several information technology projects that aim to consolidate scientific information from research domains other than the field of health behavior change. We evaluate these projects in terms of scoping (i.e., what data are collected?) and tooling (i.e., how are data collected?).

In terms of scoping, especially the biomedical field has contributed largely to this challenge. For example, SWAN [18] and Nanopublications [19] have been used to model the results or outcomes of empirical studies in the biomedical field. Moreover, SWAN is capable of recording the original hypothesis (i.e., claim) as well. However, both tools provide limited means to record contextual information on empirical studies (e.g., sample demographics, details of an experimental setup, etc.). Micropublications [20] and ECO-CollecTF [21] do constitute features to record contextual information on empirical studies. However, these tools do not distinguish different types of empirical data as separate entities. On the other hand, SALT [22] and SOLVENT [23] are two domain-independent solutions to construe knowledge graphs. SALT [22] is tailored toward capturing details of an empirical study in great depth, whereas SOLVENT [23] is particularly tailored to lightweight data capturing (at the expense of richness of data).

Subsequently, in terms of tooling, SWAN [18], Nanopublications [19], Micropublications [20], and ECO-CollecTF [21] are designed to enrich various document formats (e.g., HTML and PDF) with semantic annotations. Subsequently, these initiatives store these annotations in databases and provide digital interfaces to reason from these annotations (e.g., see [20]). Similarly, SOLVENT [23] relies on annotations to

collect data, especially via crowdsourcing, SALT [22], on the other hand, relies on annotations of LaTeX documents.

From our exploration of related work, we found that all currently available tools lack the possibility to relate empirical data to established theoretical constructs. Furthermore, only a limited number of tools supports the capturing of contextual information. However, in some fields (e.g., the field of health behavior change), this functionality is desperately needed [11, 15], to be able to make replicable predictions on the impact of a particular intervention strategy in a specific context. We aimed to overcome these limitations by developing a toolbox that empowers scientists in fields where many theories for explaining a single phenomenon exist and where consensus is not yet established on how these theories relate to each other.

Methods

This section describes the modeling steps we have undertaken to ensure that our solution used the existing concepts where appropriate. Then, it describes our evaluation strategy.

Modeling Approach

SciModeler constitutes a set of tools to encode how empirical studies support or refute one or more scientific theories, as well as record contextual information, in a knowledge representation that facilitates querying. Figure 1 details the

(interrelationships between) tools that *SciModeler* includes. The tools that are inherited from the database system are displayed in red, and the tools that are specific to *SciModeler* are displayed in blue and gray. For example, capabilities to import data, query data, and obtain query results, are native to the database system we have selected. Additionally, to import data, we have developed a web-based application to derive semantic meaning from empirical studies using PDF annotations. Other *SciModeler*-specific tools (e.g., a dedicated application to encode scientific theories, a LaTeX extension to capture semantic meaning directly from LaTeX documents, an interface to discuss mappings between empirical data and theoretical constructs among researchers, and a querying interface that can be used to query the database and display results in dedicated dashboard views) are currently still being developed.

To develop this toolbox, our first task was to derive *SciModeler*'s metamodel. The requirements for this metamodel were derived from two literature surveys. These literature surveys were focused on dissecting scientific theory, and anatomizing empirical studies, respectively, such that we gained a proper understanding of the different types of data *SciModeler* should be able to record. Subsequently, we have designed a metamodel to meet these requirements.

After we had designed our metamodel, we had to select a database model and database system to actually instantiate our proposed metamodel. Hence, to instantiate our metamodel, we have adopted a Labeled Property Graph (LPG). This graph-based approach was chosen, for its flexibility, and

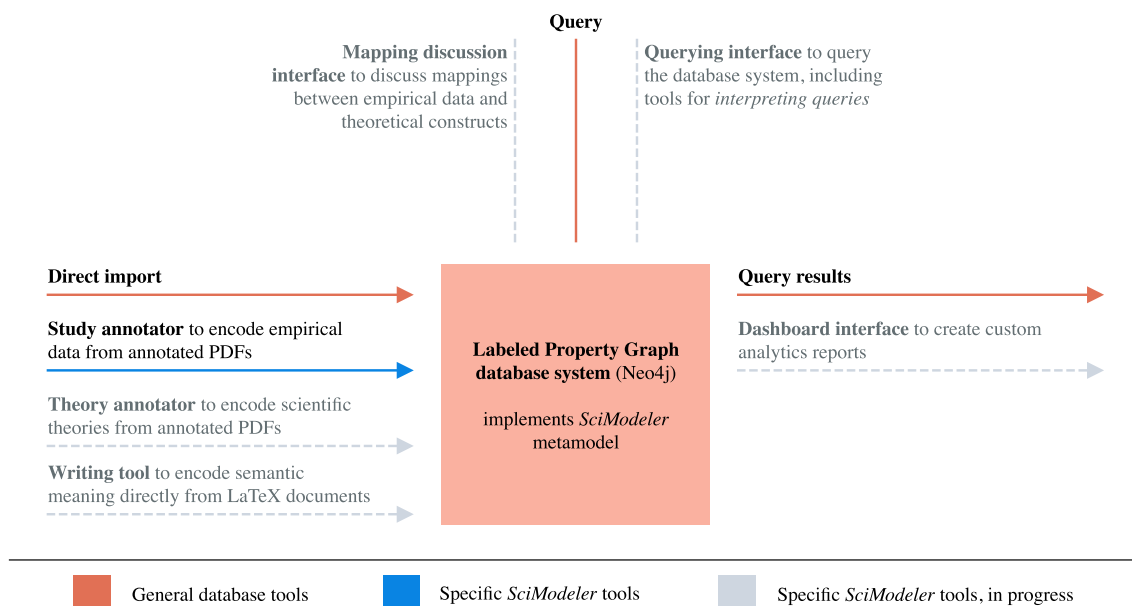


Fig. 1 Overview of tools that constitute the *SciModeler* toolbox

extensive coverage of database systems. Unlike Resource Description Framework (RDF) Triple Stores, another common approach to store and depict connected data that is often used for ontologies [15, 24], LPG allows for: (1) defining attributes on nodes, and (2) connections of the same type between the same pair of nodes [24]. Since these features were essential for implementing our metamodel without technical distractions, we have adopted LPG instead of RDF. Note, however, that LPG labels can ultimately be compiled into RDF triples [24], which may be relevant for interoperability.

Subsequently, to implement the LPG graph model, we adopted the graph database system Neo4j v4.1.3, partly because Neo4j provides extensive tools for visualizing data (e.g., by displaying query results as actual graphs), and a declarative graph querying language (i.e., Cypher [25]).

Finally, before we could evaluate the usefulness of our system in a case study, we had to design a methodology for recording data (i.e., scientific theory, as well as empirical data). First, to record a scientific theory, we proposed that a reviewer examines the original research article presenting the theory, and writes a set of import statements (i.e., using Cypher, Neo4j's graph query language) to commit the theory to the database. In general, this exercise is relatively straightforward, as scientific theories are often readily visualized as graphs, including the *constructs* and *relations* that are recorded in our database.

In contrast to capturing theories, the exercise of extracting data from articles on empirical studies is typically more challenging. Particularly, because empirical results are usually embedded in text-based documents: a format that is not easily transformed into a graph structure. Hence, to record empirical results from text-based documents, we have developed a dedicated tool: the *SciModeler* study annotator. This web-based tool derives import statements (i.e., in Cypher) from annotated PDF documents. Moreover, the study annotator permits users to annotate PDF documents directly from a web browser. Hence, users highlight text that represents a particular semantic meaning, and encode that highlighted text as an entity instance of *SciModeler*'s metamodel. Then, the user selects the appropriate attribute that the highlight represents, and associates the instance to other entity instances. The highlighted text, as well as an optional description are recorded as the attribute value. The highlighted text is recorded to reassure that the source of a piece of empirical data can easily be traced back to the original article.

Evaluation Strategy

After developing the appropriate tools, we have performed a case study to demonstrate how *SciModeler* can facilitate the consolidation of scientific knowledge in the field of health

behavior change, by facilitating: (1) recording study findings and contexts in a knowledge representation that facilitates querying, (2) mapping study outcomes with theoretical constructs to refine scientific theory, and (3) making replicable predictions on the impact of an intervention strategy in a specific context.

First, we have recorded three defying theoretical frameworks from the field of health behavior change. In the context of behavior change, theories seek to explain why, when, and how a behavior does or does not occur, and the important sources of influence to be targeted to alter the behavior [26]. Theories on behavioral change are prevalent: The book *ABC of behaviour change theories* reports 83 behavior change theories [26]; a scoping review on theories of behavior change identified 82 distinct theories [10]; and the book *Planning health promotion programs* discusses more than 40 behavior change theories [27]. From these and other sources, we have compiled a list of 103 unique behavior change theories.

In an online survey, we have requested behavioral scientists to express what theories they typically use in their behavior change initiatives. The survey was completed by 38 scientists who selected: (1) the Self-Determination Theory [28, 16 mentions], (2) the COM-B System [13, 15 mentions], and 3) the Goal-Setting Theory [29, 14 mentions] as the most useful theories of behavior change:

The COM-B System is a theory that proposes that, for a behavior to occur, an individual must have the capability (i.e., physical or psychological) and opportunity (i.e., triggered from the social or physical) to engage in the behavior, as well as the strength of motivation (i.e., “reflective” or “automatic”) to engage in it must be greater than for any competing behaviors [13]. The model emphasizes that components can interact: for example, motivation can be influenced by both opportunity and capability, which can in turn influence behavior. Behavior can then have a feedback influence upon a person's opportunity, motivation, and capability to perform the behavior again.

The Self-Determination Theory (SDT) provides a broad framework to study motivation, personality, and behaviors [30]. Central to the theory's explanation of behavior is the distinction between intrinsic motivation (i.e., motivation due to inherent interest or enjoyment) and extrinsic motivation (i.e., motivation due to external factors), and people's need for autonomy, competence, and relatedness [28, 30, 31].

The Goal-Setting Theory explains the mechanisms by which goals or intentions influence task performance [29, 32]. The theory's basic premise is that an individual's conscious ideas regulate his/her behavior (i.e., task performance). Additionally, performance can be

moderated by a number of factors including the level of commitment, the importance of the goal, levels of self-efficacy, feedback, and task complexity [29]. Furthermore, the authors model the impact of relationships between goals and their impact on satisfaction, as well as how goals act as mediators of incentives.

Additionally, we have evaluated how valuable information from three empirical studies on health behavior change could be recorded. To reliably model three exemplar empirical studies in the field of health behavior change, we drew from our own collection of empirical studies. The examples have quite diverse study setups, and are, therefore, suitable to demonstrate the expressiveness of *SciModeler*, as well as these studies provide a good basis for illustrating the usefulness of the toolbox.

Study S1 [33]. This study evaluated two design elements of an mHealth app (i.e., social proof and tangible rewards) and their impact on user engagement. It was found that the introduction of a sufficiently meaningful, unexpected, and customized extrinsic reward can engage participants significantly. During a 4-week campaign, a sample of 143 university staff members engaged in a health promotion campaign. Participants were randomly distributed over one of three treatment groups.

Study S2 [34]. This study evaluated the impact of personalized motivational messages, as compared to randomized motivational messages. It was found that personalized messages are more appreciated than random messages, but also that personalized messages do not necessarily cause a change in (long-term) behavior.

Study S3 [35]. This study evaluated social comparison as a driver of engagement with an mHealth app in preadolescents. It was found that a team-oriented environment with involvement of a natural role model is more engaging than an individually focused setting. This conclusion was drawn after a 12-week crossover experiment, including 290 preadolescent students, in which three social comparative settings were evaluated.

Finally, after we had recorded three theoretical frameworks and three empirical studies, we have explored how these theoretical frameworks and empirical studies map onto to each other, and how these relations could be represented by *SciModeler*. Moreover, we have explored how the system could be queried to consolidate knowledge. For that, we developed various queries in Cypher.

Results

In this section, we first present the results of our modeling steps, and then, we share our evaluation results.

Modeling Outcomes

Dissecting Scientific Theory

A theory comprises a set of abstract statements about reality [36]. Hence, informal explanations, unfalsifiable statements, and ideas are important, but they are not scientific theories [37, p. 23]. Instead, in a “theoretical system”, “theoretical constructs” are introduced “jointly” (i.e., associated to each other) [38, p. 32], such that a natural phenomenon and its antecedents are explained and their relations can be repeatedly tested and verified. Without theory, it is impossible to make meaningful sense of empirically generated data, and it is not possible to distinguish positive from negative results [39, p. 23].

For this study, we have assumed that a scientific *theory* comprises *constructs*, and the *relationships* between these constructs. While some related works proposed to model theories as *claims* within separate models of individual articles (e.g., SWAN [18] and Micropublications [20]), we explored a graph-based approach where theoretical elements are modeled centrally and supportive pieces of empirical data are linked to them.

Anatomy of Empirical Studies

Several frameworks for developing and reporting empirical studies have proven valuable over time (e.g., in the design of systematic literature reviews [40]). Especially, the PICO framework is commonly used in evidence-based practice (e.g., in Evidence-Based Medicine). This framework suggests that a well-defined empirical study comprises: a population, an intervention, a comparison, and an outcome. Similar frameworks were coined to be applied to different research fields. For example, the PECO framework (i.e., Population, Exposure, Comparator, and Outcomes) was designed for environmental, public, and occupational health research [41]; the SPICE framework (Setting, Perspective, Intervention, Comparison, and Evaluation) was introduced to support qualitative research [42], as well as the SPIDER framework (Sample, Phenomenon of Interest, Design, Evaluation, Research type) [40]; and finally, the ECLIPSE framework (Expectation, Client group, Location, Impact,

Professionals, ServiceE) was designed for health management research [43]. Across frameworks, the following components can be identified:

The Population refers to the community that is targeted within a study (e.g., Dutch high school students, or older adults at risk of being overweight, etc.). This concept is also referred to as the Patient group, Sample, Perspective, or Client group.

The Setting (or Location & Timing) describes when and where an intervention was evaluated [42].

The Expectation (from ECLIPSE, corresponding to the Outcome from PECO or the Evaluation of SPIDER) is the end point of interest. Once this dependent variable is known, the impact of studies addressing a similar outcome variable can be compared. Note that careful recording of this outcome variable is necessary, as a variable can sometimes be measured in different ways [44, p. 29].

The Intervention (or Phenomenon of Interest, Professionals & Service) indicates the object that is studied and that is expected to cause a difference (e.g., the administration of a medical drug) [44, p. 29].

The Comparison (or PECO's Comparator, or SPIDER's Design) is measured against the intervention. Often, the comparator is a different treatment, or alternatively the absence of a treatment.

The Impact (from ECLIPSE, corresponding to the Evaluation from SPICE) describes what results the evaluation yielded [42].

The Research type (from SPIDER) captures the study design that was adopted to evaluate the intervention [40].

Requirements Elicitation

From the dissection of scientific theory (i.e., “Dissecting Scientific Theory”), we have concluded that theories consist of constructs and relations. While previous formalisms (e.g., SWAN [18] and Micropublications [20], also see [Related Work](#)) already support the encoding of claims of individual articles, it is worth representing theories as first-class modeling concepts, which can be linked from individual studies. Regarding the coding of empirical studies, various formalisms have already been proposed. However, the systematic linking of empirical data with theoretical constructs is lacking, especially in research that involves multiple scientific domains [15]. To overcome these limitations, we propose a new metamodel that has two layers: The first layer supports the encoding of scientific theories (*ST*) and the second layer supports the encoding of empirical studies (*ES*).

For layer *ST*, we identified three information requirements (i.e., based on the section “Dissecting Scientific Theory”), aimed at representing theories as graphs:

- ST1* Record the name of the theory;
- ST2* Record the primitive constructs of the theory;
- ST3* Record the relations between these constructs.

Similarly, for layer *ES*, our synthesis of the section “Anatomy of Empirical Studies” leads to:

- ES1* Record the (characteristics of) the study population and study sample (i.e., to whom?);
- ES2* Record the setting (i.e., place and time) of the study (i.e., where, and when?);
- ES3* Record the expectation of the study (i.e., why?);
- ES4* Record the interventions and comparison treatments (i.e., what? what else?);
- ES5* Record the impact of the interventions and treatments on the study sample (i.e., how well?);
- ES6* Record the research type (i.e., how?).

While the requirements for layer *ST* and layer *ES* followed reasonably simply from the section “Dissecting Scientific Theory” and the section “Anatomy of Empirical Studies”, respectively, we found that especially the linking of the two layers is non-trivial. Regarding the interlinking of the two layers, one would ultimately like to see how specific elements of an empirical study relate to specific elements of a theory. Regrettably, many empirical studies only label interventions at the aggregate level of theories. From our modeling requirements point of view, we therefore need to support both ways of linking the empirical layer with the theoretical layer. Furthermore, concrete interventions in empirical studies can be coded differently according to one's point of view (even when aiming to minimize subjectivity). We will illustrate this challenge by means of a case study in the section “Case Study Evaluation”, but regarding modeling requirements, we conclude here that there is a need to support competing classifications and leave it up to the scientific discourse to decide which classification is the best for a specific purpose.

- ES* → *ST1* Record the relation between a theoretical construct and an actual intervention;
- ES* → *ST2* Record the argumentation for why this relation is appropriate;
- ES* → *ST3* Foster a discussion of the scientific community on a particular suggested relation.

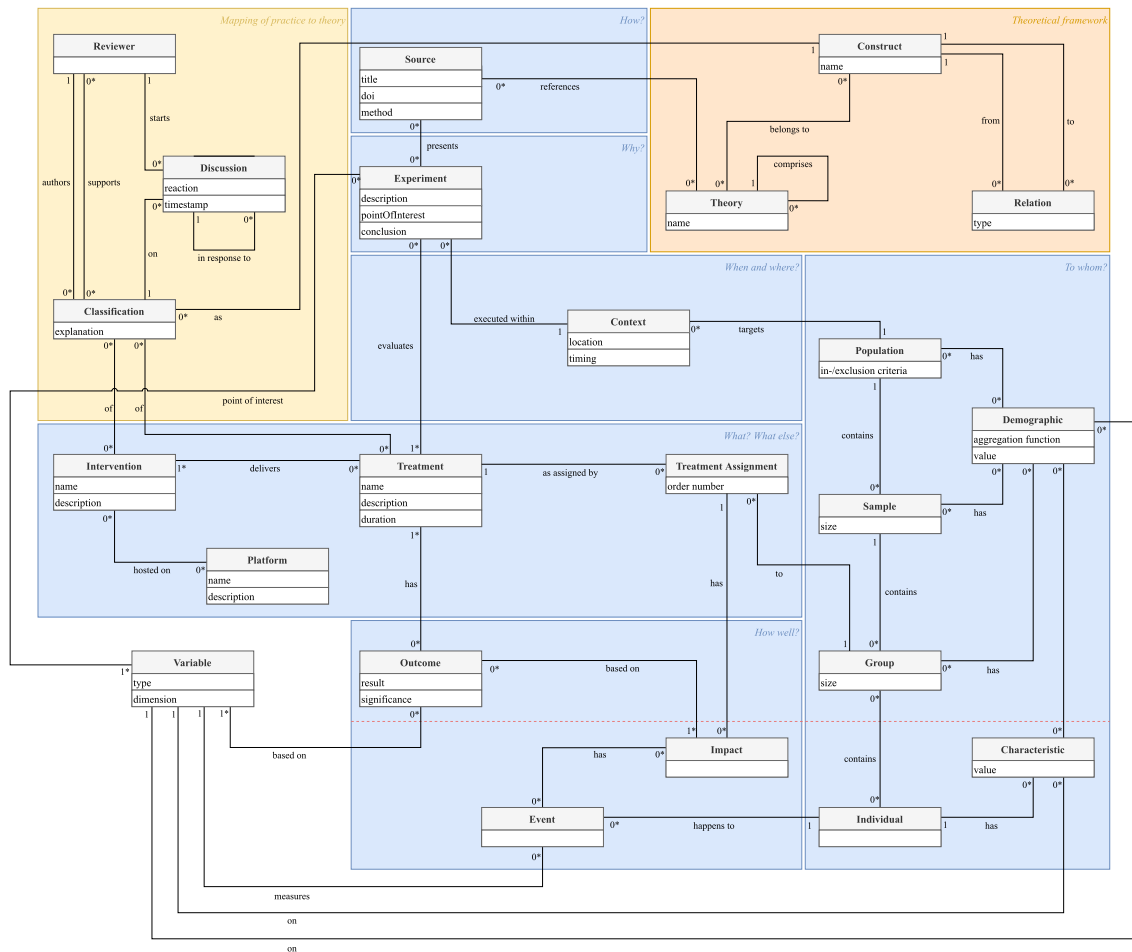


Fig. 2 SciModeler metamodel (v2.1.0) [45]

Metamodel design

Based on the modeling requirements, we have developed a metamodel that is displayed in Fig. 2. The colored rectangles in the background demonstrate what particular requirement is fulfilled by the rectangle’s enclosed entities, attributes, and associations.

The orange rectangle captures the entities, attributes, and associations that were necessary to satisfy the requirements at layer *ST*. Particularly, to: (1) record the name of a theoretical framework using the *theory* entity [ST1]; (2) record the constructs within a theoretical framework using the *construct* entity [ST2]; and record the relations between the constructs of a theoretical framework via the *relation* entity [ST3]. The relation entity has a type attribute that can have the values: *has an influence on*, *has a positive influence on*, *has a negative influence on*, *is a component of*, and *is synonym of*.

The blue rectangles depict the entities, attributes, and associations that were necessary to satisfy the requirements at layer *ES*. First, the entities *population*, *sample*, *group*, *individual*, *demographic*, and *characteristic* are necessary to record with whom a particular intervention was evaluated [ES1]. The *population* entity captures information about the audience that was targeted for a specific study. The *sample* entity records how many subjects from this population have actually participated in the study. The *group* entity distinguishes the number of participants that were exposed to a specific treatment. The *demographic* entity can be used to collect additional information about these groups on different variables. For example, this entity, its attributes and associations, may be used to record that the average age of a sample was 27. In that scenario, *age* is the dimension of the variable associated with the demographic, the aggregation function of the demographic is *average*, and the value of the demographic is 27. Note that the actual ages (i.e., recorded as *characteristics*) of the individuals within the sample may

nevertheless be undisclosed, but we may know that the average age of the sample is 27.

Second, the *context* entity is used to record where and when a study was executed [ES2]. For example, a study may be executed at a high school (i.e., location) during the winter of 2018 (i.e., timing).

Third, the *experiment* entity records the rationale behind a study [ES3]. Particularly, the point of interest, or outcome variable is recorded.

Fourth, the entities *treatment*, *treatment assignment*, *intervention*, and *platform* are used to record what treatments were assigned, and how these compare to each other [ES4]. The *intervention* entity records particularities that are present within all treatments, whereas the *treatment* entity only records particularities that are unique to a specific treatment. The *platform* entity can be used to emphasize that a set of interventions relies on shared infrastructure. For example, a marketing intervention may be administered via a phone call, and different interventions may use similar infrastructure. As an example from the software engineering domain, the Eclipse framework could be a platform on which an empirical study on plug-in development could be based. Finally, the entity *treatment assignment* can be used to assign a particular treatment to a group of participants.

Fifth, the entity *outcome* records the impact of a specific *treatment* [ES5]. Particularly, by capturing the treatment result and the significance of that result. Note that, in empirical reports, results are not often shared at the *individual* level, but rather at the *treatment* level, because the actual datasets that were obtained to derive an empirical result are typically not shared. Hence, specific information about the *characteristics* of particular *individuals*, or the *impact* the intervention has had on a particular *individual* are mostly not revealed in scientific outlets. Therefore, the entities, attributes, and associations that are displayed below the red dotted line in Fig. 2 are included for completeness, but are known to be difficult to extract from most research articles. Then again, future articles on empirical studies may cite *SciModeler* instances as online attachments that document the study setup with greater precision.

Sixth, the entity *source* is used to record the scientific article that describes the research method underpinning one or more experiments [ES6].

Finally, the yellow rectangle captures the entities, attributes, and associations that are used to map empirical data onto theoretical constructs (i.e., linking layer *ES* and layer *ST*). The *classification* entity can be used to associate (parts of) a particular intervention or treatment with a theoretical construct [ES → ST1]. Since this step relies on interpretation, an explanation from a reviewer is required [ES → ST2]. Other reviewers can support (i.e., “upvote”) a given *classification*, or commit their own [ES → ST3]. Finally, a reviewer

may start a *discussion* on a given *classification*, or start a *discussion* in response to an existing *discussion* [ES → ST3].

Case Study Evaluation

Recording Scientific Theory

To populate the database with data on selected scientific theories (e.g., the COM-B System, the SDT, and the Goal-Setting Theory), import statements (in Cypher) were manually derived from research articles on these respective theories. This task was trivial, as scientific theories are generally presented in a graph-based format, already. For example, Fig. 3a) displays the original COM-B System and Fig. 3b) displays how the constructs within the COM-B System, and their interrelationships, could be captured within *SciModeler*. Again, note that the translation of the original theoretical framework into a *SciModeler* instance was straightforward.

The import statements for recording the SDT and the Goal-Setting Theory were obtained from this procedure as well. Still, we do want to outline some particularities that were unique to these theories, and demonstrate how these particularities are handled withing *SciModeler*. First, the SDT is a meta-theory comprised of five mini-theories [30]. The notion that theories can be composed of other theories can be recorded in our system using, see Fig. 4a), the recursive relationship that the *theory* entity has with itself. Second, in the Goal-Setting Theory, the constructs “goal” and “intention” are used as synonyms. The notion of equivalent constructs can be recorded in our graph using a *relation* of type synonym; see Fig. 4b).

Recording Empirical Data

Subsequently, we have recorded data from three empirical studies. To record the relevant data, we annotated the original research articles of these studies using the *SciModeler* study annotator, see Fig. 5. After the annotations were made, this tool was used to output a set of input statements that were imported into our database. The related infrastructure—including the exemplar annotated PDF documents—is available via GitHub [46].

In the remainder of this section, we will briefly outline how some particularities for each empirical study were recorded within *SciModeler*. First, Fig. 6 displays how information on the first study’s *population*, *sample*, and treatment *groups* could be recorded in *SciModeler* (i.e., by means of the indigo, purple, and violet nodes). Additionally, *sample demographics* are recorded as well (i.e., via the pink and green nodes).

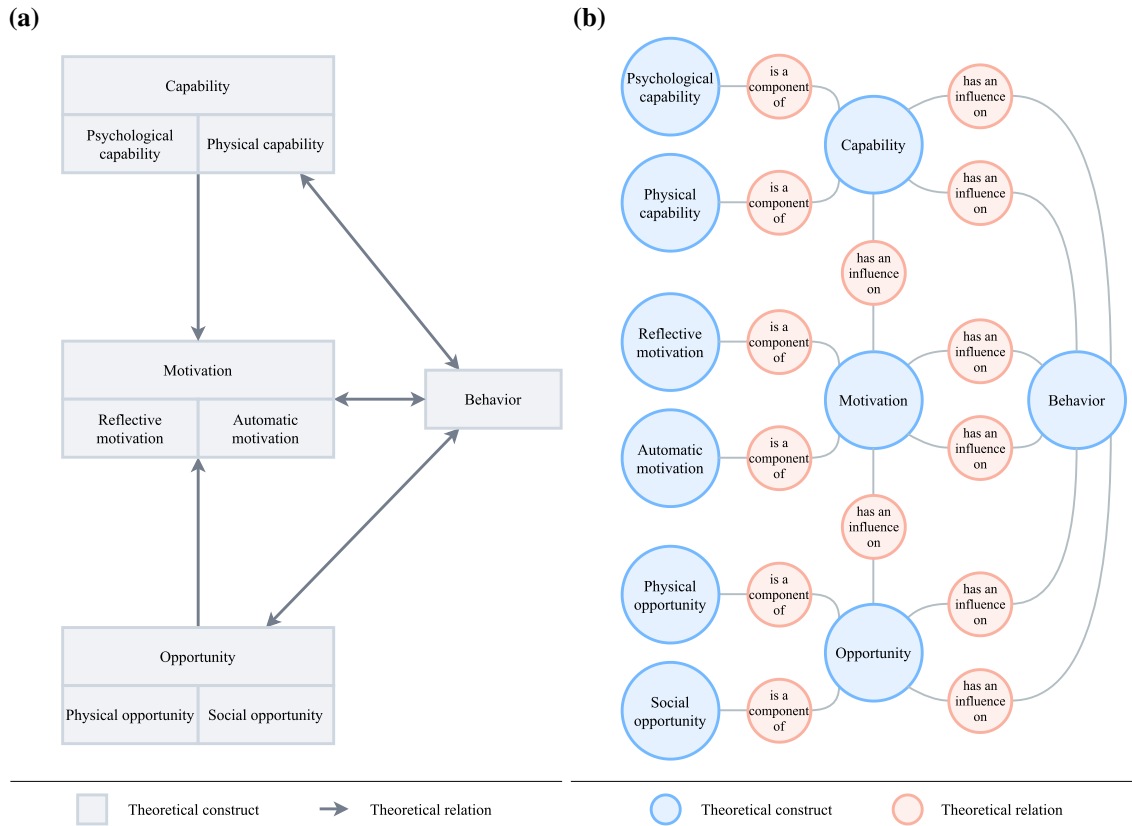


Fig. 3 **a** The original COM-B System; **b** the object diagram for the COM-B system

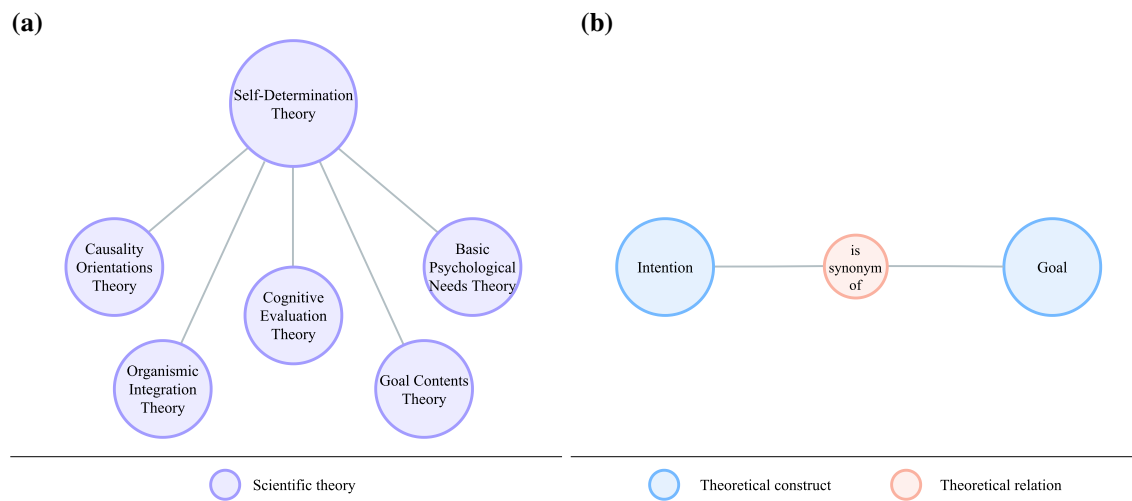


Fig. 4 **a** Object diagram that details the interrelationships between the SDT mini-theories; **b** object diagram that illustrates the use of the relation type “is synonym of”

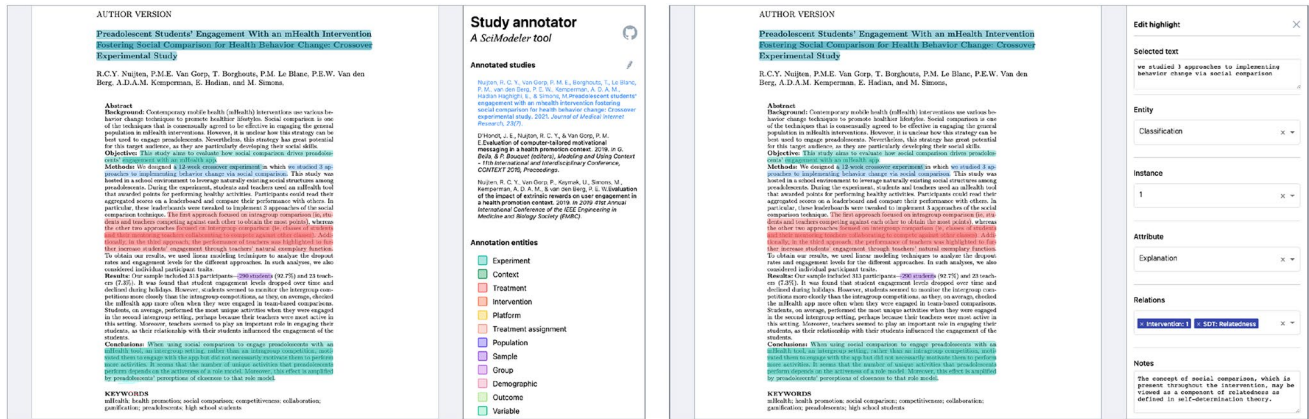


Fig. 5 The study annotator is used to generate import statements from PDF annotations



Fig. 6 Object diagram that records the population, sample and group decomposition of study S1, as well as the sample demographics

Figure 7 displays how the general *intervention* (i.e., perform health-related activities to obtain virtual points) and the two *treatments* of study S2 (i.e., personalized motivational messages, as compared to random motivational messages) can be recorded in *SciModeler* (i.e., by means of the orange and red nodes, respectively). Also, the (mHealth) platform can be recorded that is used to host the intervention and treatments (i.e., yellow node) Additionally, the *outcomes* of

the study, as well as the outcome *variables*, are recorded in the green nodes.

Study S3 employed an advanced study design. Figure 8 displays how that complex study design was captured in *SciModeler*. Particularly, in study S3, a crossover experimental design with three treatments was employed, where every treatment *group* received their treatments in 2-week periods,

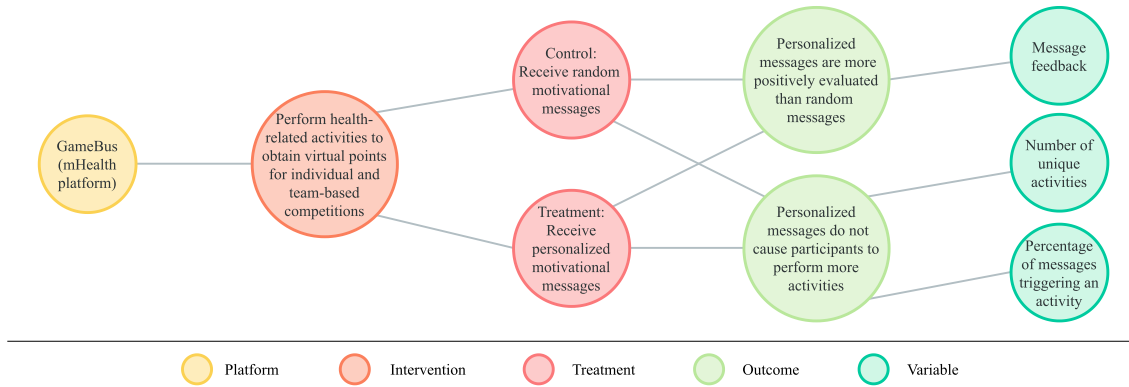


Fig. 7 Object diagram that records the intervention and treatment structure that was employed in study S2, as well as the study outcomes

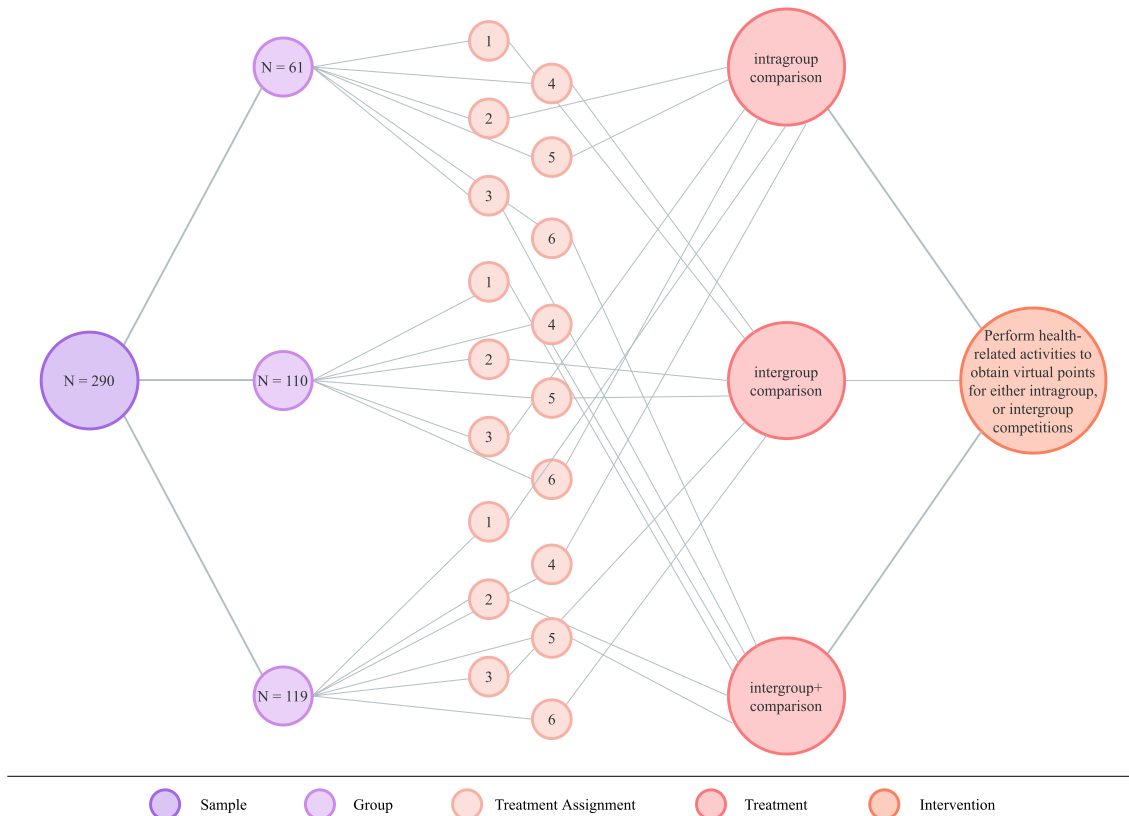


Fig. 8 Object diagram to express the complex study design of study S3

and received every treatment twice. Figure 8 displays how treatment groups (i.e., violet nodes) are linked to the treatments (i.e., red nodes) through instances of the treatment assignment entity (i.e., pinkish–orange nodes). The attribute order number on the entity treatment assignment is used to distinguish in what order the treatments were assigned to a particular treatment group.

Consolidating Scientific Knowledge

The entire *SciModeler* database can be retrieved from [47]. Subsequently, the final exercise was to link (elements of) the interventions and treatments of our empirical studies onto theoretical constructs. We mapped our studies’ interventions and treatments onto four theoretical constructs; see Fig. 9.

The interventions we had employed were similar in each empirical study (i.e., collecting virtual points for performing

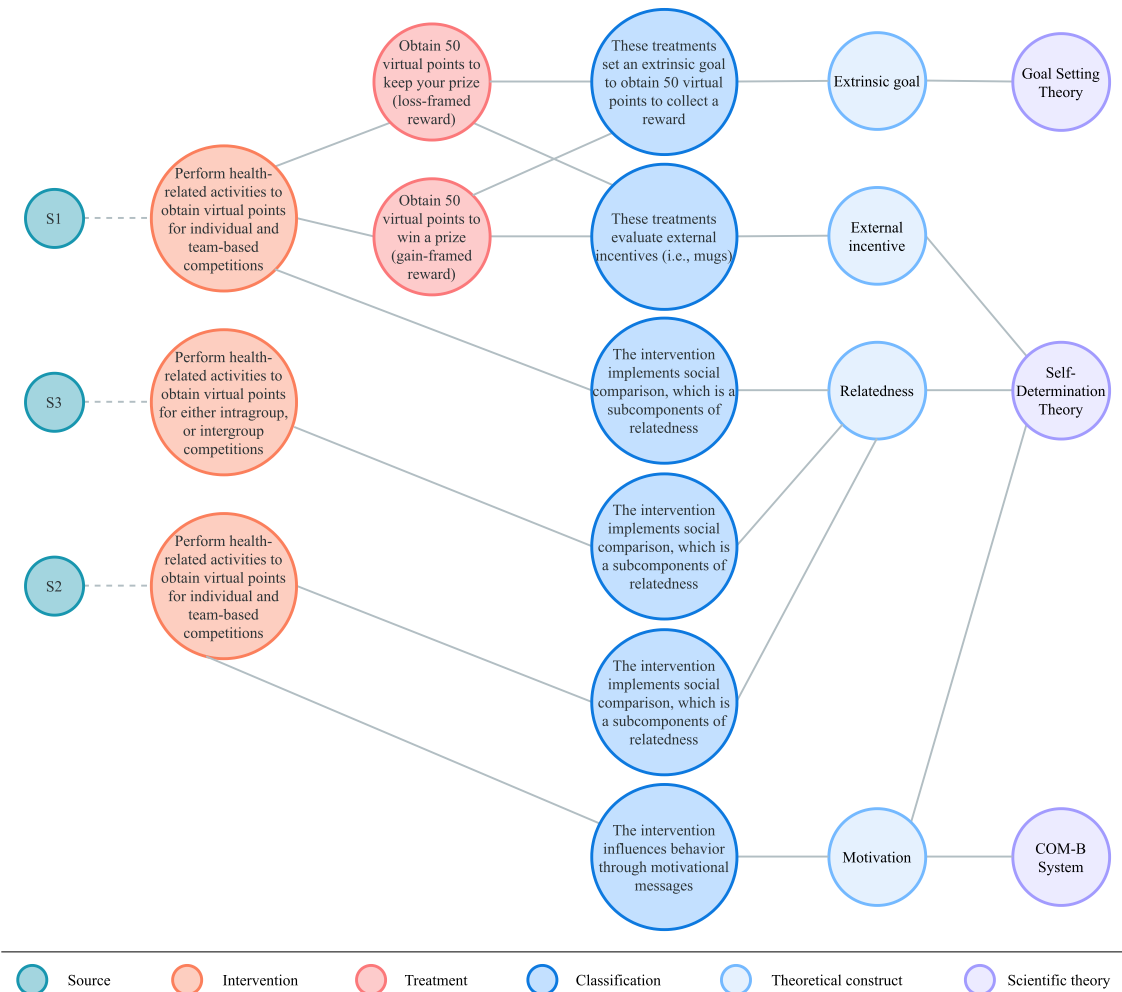


Fig. 9 Object diagram of the mapping of interventions and treatments onto theoretical constructs

health-related activities, to compare oneself to peers), and therefore related to the construct of relatedness, a concept that is expressed in the Self-Determination Theory. Furthermore, Study S1 employed tangible rewards in two of its treatments, and therefore, these treatments were mapped onto the constructs of “extrinsic goals” and “external incentives”. Finally, Study S2 employed motivational message. Hence, its intervention was also mapped onto the construct “motivation”, that both the COM-B System and the SDT include.

Finally, we could query our graph to consolidate scientific knowledge. In this section, we present three ideas for querying the graph database.

The first strategy may be adopted to refine scientific theory. One may query all interventions and treatments that address a particular theoretical *construct*. Then, one can

evaluate the outcomes these interventions and treatments had on the target variables and check whether the theory under investigation would suggest that same outcome. For our case, we may query all interventions that were associated to the construct relatedness; see Query 1a of Appendix A. We then find that there are three interventions associated with this construct, also Fig. 9. Now, we can evaluate whether the outcomes are to be expected according to our theory on relatedness, and we may update our theories accordingly. For example, one can evaluate whether suggested theoretical *outcomes* also translate to other *populations* and *contexts*. Note that a user of this system may determine herself what theoretical constructs are interesting to evaluate: one can even jointly evaluate the empirical impact of multiple constructs, if one believes several constructs represent a similar semantic meaning, see Query 1b of Appendix A. Then, one

can explore *outcomes* to evaluate whether results are similar (in a particular context), and constructs may be merged, or latent relationships may be exposed.

Second, one may want to explore promising intervention strategies for a particular context (e.g., a particular target audience). One can retrieve all experiments that target a particular *population*, or *context* to evaluate whether an *outcome* can be replicated within that *population* or *context*; see Query 2 of Appendix A.

Third, one may query all experiments that have used the same (mHealth) *platform* to evaluate whether a theoretically suggested relationship is reported consistently with (probably) similar interventions and *treatments*, in different contexts. Using Query 3 of Appendix A, one can find all *interventions* and *treatments* that were hosted using a similar *platform*. This query can also be run to explore the theoretical concepts and ideas that are implemented in a given (mHealth) platform.

Discussion

We have demonstrated the development of *SciModeler*, a toolbox for consolidating scientific knowledge in the field of health behavior change. Also, we have demonstrated the potential value of *SciModeler* by means of a case study. Even though the example queries were relatively simple, they were used to retrieve information which would be difficult to obtain reliably when only reasoning about the original research articles. We have also suggested that this basic infrastructure paves a way toward automating the simplification and merging of theories. Still, the setup in which *SciModeler* was demonstrated has various limitations, that call for future improvements.

We have found that it remains challenging to populate *SciModeler*'s database. Particularly, it remains challenging to record data from empirical studies. On the one hand, this challenge prevails, because empirical reports in the field of health behavior change are typically incomplete, or ambiguous [3]. The ambiguity of empirical reports is not unique to the field of health behavior change, as a peer review of 313 research studies found that over half (54%) of the studies did not report on the four PICO components [48]. On the other hand, populating *SciModeler*'s database remains challenging, because the data entry process currently requires (extensive) manual labor, despite the dedicated *SciModeler* study annotator we have developed. To reduce this burden, we aim to explore the possibility of a LaTeX extension to capture semantic meaning directly from LaTeX documents. Using this extension, annotations for data import could be

construed while writing scientific documents and at the same time safeguard that all relevant information is actually in the document. Then again, the availability of scientific literature in this format may be questionable as LaTeX use is historically limited to the hard sciences [49]. Hence, on the other hand, we aim to explore NLP and ML techniques for automatically mining *SciModeler* models. Regrettably, these algorithms will also suffer from the fact that many scientific publications are incomplete, or ambiguous [3, 48].

Moreover, to reduce the burden of populating *SciModeler*'s database—and in fact improve the value of the toolbox altogether—we aim to explore how the system could interoperate with other systems. For example, data could be imported from and shared among different systems. As each system has their advantages and disadvantages, in the case of *SciModeler* it may be interesting to evaluate interoperability with a system like SOLVENT [23] that is particularly tailored to lightweight data capturing. For example, SOLVENT models could be imported in *SciModeler*. Note, however, that the opposite would not be possible without information loss, as SOLVENT would not be able to consume the more extensive *SciModeler* models in their entirety. Nevertheless, this exploration paves the way for multi-step approaches, in which researchers for example: (1) generate SOLVENT models from a long list of potentially interesting research articles, (2) analyze the SOLVENT models to create a short list of actually interesting documents, (3) import the short-listed SOLVENT models into *SciModeler*, (4) complete the newly imported *SciModeler* models by adding annotations to the models' original research articles, and (5) analyze the complete *SciModeler* models to gain new insights.

Regardless of whether studies are labeled by their original authors, by other scientists, or by an artificially intelligent agent, one may want to collect community feedback on the quality of a *SciModeler* model. We have anticipated that by allowing users to discuss and “upvote” each other's *classifications* (i.e., mappings) of experimental *interventions* and *treatments* as theoretical *constructs*; however, a digital interface is not readily available yet. Additionally, future metamodel revisions should support discussion at the level of other entities and attributes too, such that the truthfulness of a particular attribute value can be measured by the degree to which reviewers agree on the information.

Finally, tooling for querying the system should be improved. A current limitation is that we do not yet provide an interface for querying the graph, other than Neo4j's built-in query explorer. To also allow possible non-expert end users to use the system, we plan to provide an interface, for instance with a set of default queries. Additionally, at the

level of the *SciModeler* metamodel, future work is to decompose the text-based node attributes into more fine-grained sub-graph structures. That would, for example, enable the query-based retrieval of studies that are recorded within the context of a high school, with a duration of at least 8 weeks per intervention. Until then, the Neo4j's query language Cypher fortunately offers support for regular expressions on node attribute values.

Conclusion

SciModeler can effectively assist scholars in: (1) refining scientific theories (e.g., by merging theoretical constructs, or exposing latent relationships), (2) exploring promising intervention strategies for a specific context (e.g., by predicting

the impact of an intervention strategy in a specific population), and (3) checking the potential impact of an intervention platform in a specific context.

As the prevalence of academic documents is growing at an exponential rate, in almost every scientific domain [20], and scientific discoveries do not span a single article anymore, but, rather span multiple articles, potentially in multiple scientific domains [15], we suggest that *SciModeler* may be applied in other domains that rely on the interpretation of empirical data (e.g., like the field of health behavior change). We invite the research community to explore with us the other domains *SciModeler* could be applied to.

Appendix A Cypher queries

Query 1a

```
MATCH (y:Construct {name: 'Relatedness'})
MATCH (y)--(z:Classification)
MATCH (z)--(i:Intervention)
MATCH (i)--(t:Treatment)
MATCH (t)--(o:Outcome)
RETURN y, z, i, t, o;
```

Query 1b

```
MATCH (y:Construct)
MATCH (y)--(z:Classification)
MATCH (z)--(i:Intervention)
MATCH (i)--(t:Treatment)
MATCH (t)--(o:Outcome)
WHERE y.name = 'Capability' OR y.name = 'Competence'
RETURN y, z, i, t, o;
```

Query 2

```
MATCH (p:Population)
MATCH (p)--(c:Context)
MATCH (c)--(e:Experiment)
MATCH (e)--(t:Treatment)
MATCH (t)--(o:Outcome)
WHERE p.inclusionExclusionCriteria =~ '(?i).*university.*'
AND c.location =~ '(?i).*Eindhoven.*'
RETURN p, c, e, t, o;
```

Query 3

```

MATCH (q:Platform)
MATCH (q)--(i:Intervention)
MATCH (i)--(t:Treatment)
MATCH (t)--(o:Outcome)
WHERE q.name =~ '(?i).*GameBus.*'
RETURN q, i, t, o;

```

Acknowledgements Parts of this work have been published and presented before at the 9th International Conference on Model-Driven Engineering and Software Development [50]. To discuss our research at this venue has benefited the quality and direction of our work.

Author Contributions Conceptualization, methodology, and validation: RN and PVG; software, data collection, data visualization, and data curation: RN; article draft preparation, RN; article review and editing: RN and PVG; funding acquisition and supervision: PVG. All authors have read and agreed to the published version of the manuscript.

Funding This work is part of the research program ‘Gamification for Overweight Prevention and Active Lifestyle’ (443001101), which is partly financed by The Netherlands Organization for Health Research and Development (ZonMw).

Data Availability The *SciModeler* Neo4j database is available via Figshare [47]. The *SciModeler* study annotator—including the exemplar annotated PDF documents—is available via GitHub (louar.github.io/SciModeler-study-annotator).

Declarations

Conflict of Interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Kirk SFL, Penney TL, McHugh T-L, Sharma AM. Effective weight management practice: a review of the lifestyle intervention evidence. *Int J Obesity*. 2011;36(2):178–85. <https://doi.org/10.1038/ijo.2011.80>.
- van Rooij T, Marsh S. eHealth: past and future perspectives. *Personaliz Med*. 2016;13(1):57–70. <https://doi.org/10.2217/pme.15.40>.
- Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, Eccles MP, Cane J, Wood CE. 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*. 2013;46(1):81–95. <https://doi.org/10.1007/s12160-013-9486-6>.
- Eckerstorfer LV, Tanzer NK, Vogrincic-Haselbacher C, Kedia G, Brohmer H, Dinslaken I, Corcoran K. Key elements of mHealth interventions to successfully increase physical activity: meta-regression. *JMIR mHealth and uHealth*. 2018;6(11):10076. <https://doi.org/10.2196/10076>.
- Sporrel K, Nibbeling N, Wang S, Ettema D, Simons M. Unraveling mobile health exercise interventions for adults: scoping review on the implementations and designs of persuasive strategies. *JMIR mHealth and uHealth*. 2021;9(1):16282. <https://doi.org/10.2196/16282>.
- Klasnja P, Hekler EB. Rethinking evaluations of mHealth systems for behavior change. *GetMobile*. 2018;22(2):11–4. <https://doi.org/10.1145/3276145.3276149>.
- Wang Y, Fadhil A, Lange J-P, Reiterer H. Integrating taxonomies into theory-based digital health interventions for behavior change: a holistic framework. *JMIR Res Protocols*. 2019;8(1):8055. <https://doi.org/10.2196/resprot.8055>.
- Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. *Health Psychol*. 2008;27(3):379–87. <https://doi.org/10.1037/0278-6133.27.3.379>.
- Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: The CALO-RE taxonomy. *Psychol Health*. 2011;26(11):1479–98. <https://doi.org/10.1080/08870446.2010.540664>.
- Davis R, Campbell R, Hildon Z, Hobbs L, Michie S. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. *Health Psychol Rev*. 2014;9(3):323–44. <https://doi.org/10.1080/17437199.2014.941722>.
- Kok G, Gottlieb NH, Peters G-JY, Mullen PD, Parcel GS, Ruiter RAC, Fernández ME, Markham C, Bartholomew LK. A taxonomy of behaviour change methods: an intervention mapping approach. *Health Psychol Rev*. 2015;10(3):297–312. <https://doi.org/10.1080/17437199.2015.1077155>.
- Fishbein M. A reasoned action approach to health promotion. *Med Decision Making*. 2008;28(6):834–44. <https://doi.org/10.1177/0272989x08326092>.
- Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci*. 2011. <https://doi.org/10.1186/1748-5908-6-42>.

14. Qin Y, Wang XR, Kang N. A meta-analysis of the overall effect of mHealth physical activity interventions for weight loss and the moderating effect of behavioral change theories, techniques, and mobile technologies. *Mob Media Commun.* 2021. <https://doi.org/10.1177/20501579211054929>.
15. Auer S, Kovtun V, Prinz M, Kasprzik A, Stocker M, Vidal ME. Towards a knowledge graph for science. *ACM.* 2018. <https://doi.org/10.1145/3227609.3227689>. <https://doi.org/10.1145%2F3227609.3227689>.
16. Tauchert C, Bender M, Mesbah N, Buxmann P. Towards an integrative approach for automated literature reviews using machine learning. In: 53rd Hawaii International Conference on System Sciences, HICSS 2020, Maui, Hawaii, USA, January 7–10, 2020. *ScholarSpace*;2020. p. 1–10. <http://hdl.handle.net/10125/63834>.
17. Fjelland R. Why general artificial intelligence will not be realized. *Human Soc Sci Commun.* 2020. <https://doi.org/10.1057/s41599-020-0494-4>.
18. Ciccarese P, Wu E, Wong G, Ocana M, Kinoshita J, Ruttenberg A, Clark T. The SWAN biomedical discourse ontology. *J Biomed Inform.* 2008;41(5):739–51. <https://doi.org/10.1016/j.jbi.2008.04.010>.
19. Groth P, Gibson A, Velterop J. The anatomy of a nanopublication. *Inf Services Use.* 2010;30(1–2):51–6. <https://doi.org/10.3233/isu-2010-0613>.
20. Clark T, Ciccarese PN, Goble CA. Micropublications: a semantic model for claims, evidence, arguments and annotations in biomedical communications. *J Biomed Semantics.* 2014;5(1):28. <https://doi.org/10.1186/2041-1480-5-28>.
21. Hobbs ET, Goralski SM, Mitchell A, Simpson A, Leka D, Kotey E, Sekira M, Munro JB, Nadendla S, Jackson R, Gonzalez-Aguirre A, Krallinger M, Giglio M, Erill I. ECO-CollecTF: a corpus of annotated evidence-based assertions in biomedical manuscripts. *Front Res Metrics Anal.* 2021. <https://doi.org/10.3389/frma.2021.674205>.
22. Groza T, Handschuh S, Möller K, Decker S. SALT—semantically annotated latex for scientific publications. *Springer.* pp 518–532. https://doi.org/10.1007/978-3-540-72667-8_37.
23. Chan J, Chang JC, Hope T, Shahaf D, Kittur A. SOLVENT. 2018;2(CSCW):1–21. <https://doi.org/10.1145/3274300>.
24. Besta M, Peter E, Gerstenberger R, Fischer M, Podstawski M, Barthels C, Alonso G, Hoefler T. Demystifying graph databases: analysis and taxonomy of data organization, system designs, and graph queries. *arXiv preprint.* 2019. [arXiv:1910.09017](https://arxiv.org/abs/1910.09017).
25. Francis N, Green A, Guagliardo P, Libkin L, Lindaaker T, Marsault V, Plantikow S, Rydberg M, Selmer P, Taylor A. Cypher: an evolving query language for property graphs. In: *SIGMOD '18: Proceedings of the 2018 International Conference on Management of Data.* ACM. 2018. <https://doi.org/10.1145/3183713.3190657>.
26. Michie S, West R, Campbell R, Brown J, Gainforth H. *ABC of behaviour change theories.* Silverback Publishing; 2014.
27. Eldredge LKB, Markham CM, Ruiter RA, Fernández ME, Kok G, Parcel GS. *Planning health promotion programs: an intervention mapping approach.* Hoboken: Wiley; 2016.
28. Deci EL, Ryan RM. *Intrinsic motivation and self-determination in human behavior.* New York: Plenum Publishing Co; 1985.
29. Locke EA, Latham GP. Building a practically useful theory of goal setting and task motivation: a 35-year odyssey. *Am Psychol.* 2002;57(9):705–17. <https://doi.org/10.1037/0003-066x.57.9.705>.
30. Ryan R. Self determination theory and wellbeing. *Soc Psychol.* 2009;84(822):848.
31. Deci EL, Ryan RM. Facilitating optimal motivation and psychological well-being across life's domains. *Canadian Psychol/Psychologie canadienne.* 2008;49(1):14–23. <https://doi.org/10.1037/0708-5591.49.1.14>.
32. Locke EA. Toward a theory of task motivation and incentives. *Organiz Behav Hum Perform.* 1968;3(2):157–89. [https://doi.org/10.1016/0030-5073\(68\)90004-4](https://doi.org/10.1016/0030-5073(68)90004-4).
33. Nuijten RCY, Van Gorp PME, Kaymak U, Simons M, Kemperman ADAM, den Berg PEWV. Evaluation of the impact of extrinsic rewards on user engagement in a health promotion context. *IEEE.* 2019. <https://doi.org/10.1109/embc.2019.8856296>.
34. d'Hondt JE, Nuijten RCY, Van Gorp PME. Evaluation of computer-tailored motivational messaging in a health promotion context. *Springer.* 2019. p. 120–133. https://doi.org/10.1007/978-3-030-34974-5_11.
35. Nuijten RCY, Van Gorp P, Borghouts T, Blanc PL, den Berg PV, Kemperman A, Hadian E, Simons M. Preadolescent students' engagement with an mHealth intervention fostering social comparison for health behavior change: Crossover experimental study. *J Med Internet Res.* 2021;23(7):21202. <https://doi.org/10.2196/21202>.
36. Reynolds P. *A primer in theory construction.* Indianapolis: Bobbs-Merrill; 1971.
37. Popper K. *The logic of scientific discovery.* Routledge; 1959.
38. Hempel CG. *Fundamentals of concept formation in empirical science.* University of Chicago Press; 1952.
39. Kerlinger FN. *Foundations of behavioral research.* New York. 1986.
40. Cooke A, Smith D, Booth A. Beyond PICO. *Qualitative Health Res.* 2012;22(10):1435–43. <https://doi.org/10.1177/1049732312452938>.
41. Morgan RL, Whaley P, Thayer KA, Schünemann HJ. Identifying the PECO: a framework for formulating good questions to explore the association of environmental and other exposures with health outcomes. *Environ Int.* 2018;121:1027–31. <https://doi.org/10.1016/j.envint.2018.07.015>.
42. Booth A. Clear and present questions: formulating questions for evidence based practice. *Library hi tech.* 2006;24(3):355–68. <https://doi.org/10.1108/07378830610692127>.
43. Wildridge V, Bell L. How CLIP became ECLIPSE: a mnemonic to assist in searching for health policy/management information. *Health Inf Libraries J.* 2002;19(2):113–5. <https://doi.org/10.1046/j.1471-1842.2002.00378.x>.
44. Melnyk BM, Fineout-Overholt E. *Evidence-based practice in nursing and healthcare: a guide to best practice.* Lippincott Williams & Wilkins; 2005.
45. Nuijten RCY, Van Gorp PME. SciModeler metamodel. *Figshare.* 2020. <https://doi.org/10.6084/m9.figshare.13347275>.
46. Nuijten RCY, Van Gorp PME. Louar/SciModeler-study-annotator: v1.0.0. *Zenodo.* 2022. <https://doi.org/10.5281/ZENODO.6341664>. <https://zenodo.org/record/6341664>.
47. Nuijten RCY, Van Gorp PME. SciModeler database. *Figshare.* 2020. <https://doi.org/10.6084/m9.figshare.13160141>.
48. Thabane L, Thomas T, Ye C, Paul J. Posing the research question: not so simple. *Health Inf Libraries J.* 2008;56(1):71–9. <https://doi.org/10.1007/s12630-008-9007-4>.
49. Martínez-López JI, Barrón-González S, López AM. Which are the tools available for scholars? a review of assisting software for authors during peer reviewing process. *Publications.* 2019;7(3):59. <https://doi.org/10.3390/publications7030059>.
50. Nuijten RCY, Van Gorp P. SciModeler: A metamodel and graph database for consolidating scientific knowledge by linking empirical data with theoretical constructs. In: *Proceedings of the 9th International Conference on Model-Driven Engineering and Software Development. SCITEPRESS - Science and Technology Publications.* 2021. <https://doi.org/10.5220/0010315503140321>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.