

A scoping review of ontologies related to human behaviour change

Emma Norris*¹

¹ Centre for Behaviour Change, University College London, WC1E 6HB, UK

Ailbhe N. Finnerty¹

Janna Hastings²

² Department of Biological Sciences and Babraham Institute, University of Cambridge, CB22 3AT, UK

Gillian Stokes³

³ EPPI-Centre, University College London, WC1H 6XX, UK

Susan Michie¹

Key words: ontology, behaviour change, quality assessment, scoping review

Correspondence to: Emma Norris; emma.norris@ucl.ac.uk

Abstract

Ontologies are classification systems specifying entities, definitions and inter-relationships for a given domain, with the potential to advance knowledge about human behaviour change. A scoping review was conducted to: i) identify what ontologies exist related to human behaviour change, ii) describe the methods used to develop these ontologies and iii) assess the quality of identified ontologies. Using a systematic search, 2303 papers were identified. Fifteen ontologies met the eligibility criteria for inclusion, developed in areas such as cognition, mental disease and emotions. Methods used for developing the ontologies were expert consultation, data-driven techniques and re-use of terms from existing taxonomies, terminologies and ontologies. Best practices used in ontology development and maintenance were documented. The review did not identify any ontologies representing the breadth and detail of human behaviour change. This suggests that advancing behavioural science would benefit from the development of a behaviour change intervention ontology.

Introduction

Behaviour change interventions are coordinated sets of activities that are designed to change specified behaviour patterns¹. While there are many examples of behaviour change interventions, as evidenced by numerous systematic reviews in the Cochrane Library², their effects are highly variable³. Such behaviour change interventions are also very heterogeneous in the techniques they employ, their settings, populations, delivery methods and target behaviours³ and the theories used to inform them^{4,5}. The accelerating volume of published behaviour change intervention evaluations makes the synthesis of this evidence increasingly difficult for researchers, practitioners and policy-makers to manage⁶. In order to improve behaviour change interventions as rapidly and efficiently as possible, it is essential that the value of published behaviour change research is maximised⁷.

A method that has helped to unify the understanding and reporting of behaviour change interventions has been the application of classification techniques; these have also been used productively in the natural and biosciences⁸. For example, typologies (classifications of general types) have been developed to classify micro-level aspects of the physical environment related to behaviours, as in the TIPPMME typology⁹. Taxonomies (classifications into ordered categories)⁸ have been used to group and define key aspects of behaviour change. The ‘content’ of interventions (their putative active components) has been taxonomically expressed via Michie’s taxonomy of behaviour change techniques (BCTTv1) with 93 BCTs organised into 16 groupings¹⁰. Similar taxonomies to represent the mode of delivery¹¹ and mechanisms of action¹² of interventions are under development. Such classification systems allow similarities and differences between interventions to be made explicit; these have been very useful for evidence syntheses^{13,14}.

Ontologies extend the hierarchical nature of taxonomies by providing a set of: i) unique and unambiguous identifiers representing types of entity (objects, attributes and/or processes), ii) labels and definitions corresponding to these identifiers, and iii) specified relationships between the entities¹⁵⁻¹⁷ (Box 1). Ontologies allow for more detailed knowledge to be expressed than via taxonomies. For example, whereas a taxonomy could express a parent-child relationship such as

Scoping review of ontologies for behaviour change

“intention is_a mechanism of action”, an ontology allows for more complex and multiple relationships such as “intention is_part_of mechanism of action ‘and ‘is_measured_by self-report’¹⁸. Sophisticated logical axioms can be used to define computationally aspects of domain knowledge within ontologies, enabling downstream computational reasoning. Ontologies also allow different theoretical perspectives and conceptual frameworks to be integrated within a single field of study, whilst also facilitating the comparisons of concepts across fields^{17,18}. They furthermore allow large datasets to be synthesized efficiently using computation^{17,18}. Ontologies are manually edited and updated according to new developments in the given domain¹⁵, providing an active basis for shared knowledge and understanding.

Ontologies have been used to revolutionise knowledge in large domains. For example, a global, ongoing effort to provide an up-to-date, comprehensive, computational model of biological systems is evident in the Gene Ontology <http://www.geneontology.org/>¹⁹. This provides a regularly updated and openly searchable source of knowledge about the biology of genes across all species using agreed upon entities (biological processes, functions and cellular components) and relationships (e.g regulation) within the growing field of molecular biology. The Gene Ontology has been used to annotate more than 100,000 peer reviewed scientific publications, providing an aggregate knowledge base that would not be possible without the ontology¹⁸. Using enrichment analysis, the Gene Ontology allows continual new insights to be derived for novel gene sets²⁰. Computation of knowledge using ontologies also facilitates evidence synthesis, allowing systematic searches to be automatically and continually updated²¹. Repositories such as BioPortal²² have been established to make ontologies freely downloadable to users. A key example of such a repository is the Open Biological and Biomedical Ontology (OBO) Foundry: a portal for the hosting and maintenance of ‘gold standard’ ontologies²³ adhering to clearly defined principles of best practice²⁴.

-----Insert Box 1 around here -----

Given the impact of ontologies in other domains, the potential of ontologies to facilitate knowledge synthesis in behaviour change is being investigated in the Human Behaviour-Change Project (www.humanbehaviourchange.org). This collaboration between behavioural scientists, computer scientists and systems architects aims to provide tools to enable researchers, practitioners and policy-makers to address the ‘big question’ of behaviour change: ‘What works, compared with what, how well, with what exposure, with what behaviours, for how long, for whom, in what settings and why?’⁶ This paper reports a scoping review conducted as part of this project, aiming to: i) identify ontologies related to human behaviour change that could be re-used in part or in full in the development of an ontology for the domain of human behaviour change, ii) describe the methods used to develop these ontologies, and iii) assess the quality of the content of the identified ontologies with respect to their potential for re-use.

Results

2427 records were identified via electronic databases (k=2415), ontology repositories (k=11) and ontology experts (k=1). After removing duplicates, 2303 records remained for screening. 2249 records were excluded for not addressing the research question at the title and abstract review stage. After reviewing full-texts, reasons for exclusion were irrelevance to the topic of this review (k=11), no description of the ontology’s development (k=14), ontologies not being fully accessible online or downloadable following contact with authors (k=13) and not being in English (k=1). Fifteen records met the inclusion criteria (Figure 1).

-----Insert Figure 1 around here -----

Existing ontologies related to human behaviour change

Descriptions of the identified ontologies are provided in Table 1, with a brief description of each ontology available in Supplementary Notes 1 and graphs depicting related measures provided in Supplementary Figure 1. Ontologies represented areas of mental processes and cognitions ^{25, 26, 27},

mental^{28,29} and physical disease^{30,31,32}, psychological experimental design³³, emotions^{34,35}, epidemiology³⁶ and healthcare^{37,38}. Identified ontologies were typically of medium scale, featuring between 100-1000 classes (entities; Table 1), with only the OBO Foundry-approved Human Disease Ontology (DOID)³⁰ having >10,000 classes (entities). Eleven different expressivity codes were found across the 15 ontologies (codes and descriptions of these are provided in Table 1). The majority of identified ontologies were accessible via OBO Foundry²³ (9/15) or BioPortal²² (11/15) repositories, with others made available via project websites^{34,39} or GitHub²⁹ (online software development platform).

-----Insert Table 1 around here -----

Many terms of relevance to human behaviour change were identified from these ontologies, with examples from each ontology provided in Supplementary Notes 1. Relevant terms included describing populations that may be targeted in interventions, primarily identified in the Ontology of Medically Related Social Entities (OMRSE)³⁷, such as ‘Smoker role’ (OMRSE_00000039) defined as ‘a role that inheres in an organism and is realized by habitually smoking tobacco products’. Classes for cognitions targeted as mechanisms of action within behaviour change were identified, such as ‘Positive emotion’ (MFOEM_000211) defined as ‘an emotion that has a positive valence’³⁵. A large range of symptoms and outcomes used as markers of behaviour change were also classified, such as ‘Alcohol consumption’ (NBO_0000131) defined as ‘a drinking behaviour associated with the intake of alcohol’²⁷ and ‘Behavioural symptom’ (SYMP_0019182) defined as ‘a nervous system symptom that involves, is related to, or emphasizes behaviour’³¹.

The ‘Health Behaviour Change Ontology’ (HBCO) of Bickmore and colleagues³⁹ was most closely related to human behaviour change (Table 1). This ontology was developed for use in a specific context: an automated behaviour change counselling system to simulate dialogue between a human health counsellor and clients. It was based around the Transtheoretical Model of behaviour change⁴⁰

to improve exercise levels and fruit and vegetable consumption. Possibly because of the specificity of its context, HBCO provides incomplete representations of behaviour change theories and concepts. It focuses on the Transtheoretical Model only and does not accurately capture the stage transitions indicated within it. As shown in a screenshot of the ontology file in Figure 2, ‘Termination’, one of the six stages of change posited by the theory⁴⁰ is missing. Stages are presented in an order not commensurate with the theory (should range from Pre-Contemplation to Termination) and relationships between stages are not indicated. Other key theories⁴, concepts of behaviour change such as behaviour change techniques¹⁰ and recommended reporting aspects such as TIDieR⁴¹ are not incorporated. It was therefore concluded that none of the ontologies we identified captures the full breadth and detail required to adequately describe and explain human behaviour change interventions.

-----Insert Figure 2 around here -----

Methods used to develop these ontologies

A summary of the methods used in the identified ontologies is shown in Table 2. Many ontologies explicitly linked their entities to other taxonomies (3/15), terminologies (6/15) or ontologies (14/15). For example, the Mental Disease Ontology (MFOMD)²⁸ incorporates entities from the DSM-V⁴² terminology to capture psychological disorders. Ten out of 15 ontologies were structured using the Basic Formal Ontology (BFO)¹⁵, a small, upper level ontology which provides a structure for development by dividing entities into two categories of ‘continuants’ (representing objects and spatial regions) and ‘occurents’ (representing processes extending over time; Box 1). Upper level structuring of ontologies using BFO is strongly encouraged for all ontologies registered on OBO Foundry⁴³. Identified ontologies also incorporated classes from other ontologies to integrate terms based on domain areas other than the primary focus of the ontology. For example, the Cognitive Paradigm Ontology (CogPO)³³ incorporates many terms from the Information Artifact Ontology⁴⁴ to describe study materials in cognitive experimentation, such as ‘action specification’ (IAO_0000007) defined as ‘a directive information entity that describes an action the bearer will take’.

User feedback was described for six out of 15 identified ontologies. This typically occurred after the ontology was launched online by inviting public comments through an online portal such as a website or Wiki page^{25,33,37}. Others reported using expert consultation during ontology development, prior to launch^{30,32,39}. For example, the Exposure Ontology (ExO)³² used a four-phase development process: after initial development from literature reviews and testing, a working group of experts critiqued the draft ontology. The final phase was seeking public feedback on the published ontology for further refinement, via repositories such as OBO Foundry and BioPortal.

Four out of 15 ontologies used data-driven techniques to identify classes and relationships.

Techniques included annotating published literature during ontology development, that is, marking of documents to identify ontological terms^{27,32}. Data from health social network datasets³⁸ and external devices³⁴ were also used in ontology development.

-----Insert Table 2 around here -----

Quality assessment of ontologies

A summary of the quality assessment used for the identified ontologies is shown in Table 2. Eleven out of 15 ontologies provided Uniform Resource Identifiers (URIs) for each class, such as ‘MF_0000027’ representing the class ‘Planning’ in the Mental Functioning (MF) ontology²⁶. Nine out of 15 ontologies had clear natural language definitions for all terms (i.e non-overlapping terms which lack redundancy²⁴), such as ‘Alcohol dependence’ (DOID_0050741), defined as ‘a substance addiction in which the substance that is compulsively consumed is alcohol’³⁰. Ontologies were assessed as not having clear definitions if entities were mostly undefined³¹, or if definitions were extremely lengthy³⁶. Graphs depicting measures of these quality assessments are provided in Supplementary Figure 1.

Twelve out of 15 ontologies were assessed as having a well-organised, hierarchical structure, including those developed around BFO such as the Exposure Ontology³² and Ontology of Medically Related Social Entities³⁷. Ontologies assessed as having unclear structures tended to lack specified inter-relationships between entity categories²⁵. Thirteen out of 15 ontologies were found to be logically consistent, i.e. contain no contradictions in their axioms, which is an important quality indicator and pre-requisite for their use in automated reasoning applications. Two out of 15 ontologies were shown to have been evaluated in their accompanying papers, via incoming data from a tangible user interface known as an ‘Emoti-Picture Frame’³⁴ and via evaluation metrics of incoming patient data³⁸ (more information in Supplementary Notes 1). Finally, 11 out of 15 ontologies were assessed as ‘maintained’, that is, updated to reflect changes in scientific consensus over time. Maintenance was typically performed via ad-hoc updates, issue trackers and online feedback portals. The Health Behaviour Change Ontology³⁹ has not been updated or maintained since its development in 2011, preventing interested parties contributing to its content.

Discussion

Fifteen ontologies related to human behaviour change were identified in this scoping review. These ontologies captured domains broadly related to behavioural science, with various entities identified related to human behaviour change. None captured the breadth and detail required to adequately describe and understand human behaviour and how to change it.

All but one ontology had explicit inter-connections with other ontologies, reflecting a key principle of inter-operability between groups of developers²⁴. These collaborations between resources highlight a primary principle of ontologies, which is to reuse existing classification systems if definitions and structures are appropriate to the domain in question^{24,45}. Incorporating user feedback to refine ontologies throughout the development process and after launch was relatively common, but the use of data-driven techniques such as manual annotations as input into ontology development, was

described less frequently. Higher quality ontologies included URIs, clear definitions and structures, and an explicit strategy for ontology maintenance.

These ontologies demonstrate global ongoing work to systematise the accumulation and computation of knowledge in areas related to human behaviour change. Ontologies exist across a broad range of disciplines and domains, with those covering the domain of human behaviour largely exploring mental health, cognition and healthcare. The large number of quality criteria met by these ontologies may reflect the criteria for inclusion in the OBO Foundry (9/15). Adherence to clear principles of ontology development and maintenance are required for an ontology to qualify for approval by them²⁴.

The one example we identified of an ontology explicitly within the domain of behaviour change was the Health Behaviour Change Ontology³⁹. This was shown to be largely based around the Transtheoretical Model⁴⁰, covering a restricted range of behaviour change theoretical concepts⁴. Relationships between stages of change in the Transtheoretical Model were absent, giving an incomplete picture of the suggested pathways to behaviour change that it proposes. The ontology was not maintained following its publication in 2011, meaning that any suggested changes to the ontology have not been incorporated and collaborations are unlikely. The lack of maintained ontologies in the area of behaviour change is in contrast to other fields which have been revolutionised by the implementation of continually updated ontologies, such as the Gene Ontology^{3,19}. There is evidently much work to be done to have behaviour change represented in the evolving space of ontologies. Development and maintenance of a behaviour change ontology would allow knowledge in this diverse research area to be efficiently shared and built. Behaviour change terms and classifications could then be incorporated into other ontologies, further broadening the awareness and reach of the collective work in the field of behavioural science.

Several ‘best practices’ for ontology development were identified in reviewing the methods and assessing the quality of the ontologies in this review. First, the review’s inclusion criteria required

ontologies to be downloadable in full, meaning that 14 potentially relevant ontologies were not included (Figure 1). For example, ONTOPSYCHIA⁴⁶, describing social and environmental determinants for psychiatry, could not be accessed as authors restricted downloads of the ontology file on BioPortal and elsewhere at the time of this review, and the authors did not respond to personal communication requesting access. This highlights accessibility as a key issue for ontologies, highlighted in OBO Foundry principles²⁴ and also in the FAIR principles for data sharing and reuse more generally⁴⁷. To ensure that ontologies provide open knowledge in any given domain, they must be publicly available and freely downloadable, via portals such as OBO Foundry and BioPortal⁶.

Secondly, the majority of identified ontologies had explicit links to other resources, mostly to other ontologies. This reflects OBO Foundry principles of reusing existing terms from other ontologies if definitions show them to be suitable²⁴. Reusing terms across ontologies allows for inter-connections across disciplines and communities, as well as limiting duplication of similar terms across the global network of ontologies. The Minimum Information to Reference an External Ontology Term (MIREOT) guidance⁴⁵ provides ontology developers with guidance on the information needed from a developing (target) and a developed (source) ontology to allow successful merging of terms across classification structures. A behaviour change ontology will likely require links to many existing ontologies, including those identified in this paper such as the Ontology of Medically Related Social Entities³⁷ to capture demographic aspects of the intervention, population and also healthcare settings.

Thirdly, as indicated by ontologies involving user feedback in their development, it is essential that a behaviour change ontology provides an open invitation for users to comment on the addition of new terms and adjustment of existing terms and relationships within it⁶. To enable this, the placement and design of a feedback portal should be implemented with the respondent audience in mind. To maximise feedback from a diverse range of interested parties, the terminology and technology used to facilitate feedback should not require understanding of ontological terminology and coding language. This will require co-design and user testing of the ontology feedback portal⁴⁸. To ensure success of the ontology in driving forward research in behavioural science and to facilitate collaboration across

Scoping review of ontologies for behaviour change

other ontologies, it is essential that development and maintenance of a behaviour change ontology should follow these best practices proposed by ontologists^{24,45}. We will release the methodology for development of the BCIO via the Open Science Framework. This will include information about our stakeholder involvement in ontology development. Methods for co-designing the development of the online platform will be added at a later point in the project.

Strengths and limitations of review

This scoping review was conducted systematically^{49,53} including publication of a review protocol²⁷, incorporating expert feedback and quality assessment and following PRISMA-ScR guidance for scoping review reporting⁵². Unlike longer-standing reporting guidelines in the health sciences, for example those from the Equator network⁵⁰, recommendations for reporting of ontology development have only recently been published, via the ‘Minimum Information for Reporting an Ontology (MIRO) guidance⁵¹. As this guidance came after the development of the ontologies we identified, it is possible that aspects of development were not reported and are thus missing from our review. Quality assessment of whether ontologies were evaluated was restricted to the identified papers only, without also searching additional papers citing the given ontologies. It is likely that subsequent papers by the ontology authors or others may have evaluated the ontologies, which will not be captured here.

Future Research

This review has identified a number of ontologies, entities, methods and best practices in developing ontologies that have the potential to inform the understanding of behaviour and behaviour change. However, none meet the criteria for an ontology of human behaviour or of behaviour change interventions, meaning there is scope for developing such an ontology.

Findings from this scoping review are being incorporated into the ongoing Human Behaviour-Change Project www.humanbehaviourchange.org, which is developing a Behaviour Change Intervention Ontology (BCIO)⁶. The BCIO consists of entities that are key to answering questions about behaviour

Scoping review of ontologies for behaviour change

change intervention effectiveness. It comprises of two upper-levels entities: ‘BCI scenario’ (a sequence of events consisting of a behaviour change intervention, its target behaviours, mechanisms of action and contextual factors influencing the outcome of the intervention) and ‘BCI evaluation’ (a comparison between two or more behaviour change intervention scenarios)⁶. The BCIO will be continually evaluated and updated based on new scientific developments and user feedback for as long as funding can be secured. Long-term models for ontology maintenance need to be developed.

As seen in data-driven methods used by some ontologies in this review, manual annotations by the HBCP’s researchers based on the BCIO are being used to direct the ontology’s development. These annotations are also being used to train an automated feature extraction system, working towards automatic identification of behaviour change entities and their relationships in published scientific literature via machine learning and reasoning algorithms. These extracted data will be used to populate a database which will be freely accessible via a user interface, allowing policy-makers, practitioners and researchers to query the evidence based on human behaviour change⁶. Automated and intelligent synthesis of behaviour change intervention knowledge using the BCIO will allow nuanced querying of up-to-date evidence²¹. Stakeholder input on the content of BCIO will also be sought throughout from potential users including researchers, policy-makers and practitioners⁶. The Human Behaviour-Change Project aims to revolutionise our ability to synthesise, interpret and generate new hypotheses and evidence about behaviour change, via the application of a specifically developed ontology.

Conclusions

This scoping review has identified, summarised and critiqued ontologies related to human behaviour change. Best practices to incorporate in ontology development were identified; this included providing an easily accessible ontology file, linking to existing ontologies and incorporating user feedback into ontology development and maintenance.

No ontology capturing the breadth and detail of behaviour change was identified, showing room for development in this area, including the development of a behaviour change intervention ontology.

Methods

This scoping review is reported according to PRISMA-ScR reporting criteria for scoping reviews⁵².

The review was pre-registered, with its protocol published on Prospero⁵³ (CRD42017079990).

Identification of ontologies

Three search methods were used to identify ontologies related to human behaviour change during October-November 2017.

1) *Electronic databases*: In order to capture publications across behavioural, health, computer and information sciences databases, we searched the following eight databases frequently used within these sciences: ACM Digital Library, Google Scholar (first 15 pages), PubMed, PsycINFO, JSTOR, Web of Science, Scopus and IEEE Xplore. Abstracts and titles were searched with three separate strings representing: 1) ontologies, 2) humans, and 3) behaviour, with only papers published since 1997 searched to reflect the period of ontology publication. The final search strategy for PubMed can be found in Supplementary Notes 2. References of included papers were also searched.

2) *Key ontology sources*: Repositories (OBO Foundry²³ and BioPortal²²) were manually searched to identify published ontologies, with relevant ontologies under development (Cochrane's PICO Ontology <http://linkeddata.cochrane.org/pico-ontology> and medical vocabularies (SNOMED⁵⁴ and MedDRA⁵⁵) also searched for relevant content.

3) *Expert consultation*: Following 1 & 2, an ontology expert (J.H), with extensive experience of developing OBO Foundry ontologies, identified any additional relevant ontologies.

Inclusion and exclusion criteria

Records were included in the review if they represented ontologies of topics related to aspects of human behaviour change, such as psychology, health and epidemiological domains. Records were excluded if they presented applications or reviews of existing ontologies without describing the ontology's development, reported ontologies that were not fully accessible online or downloadable in a machine-readable ontology representation format such as OWL, following direct contact with the authors, or were not published in English.

Data extraction

Titles and abstracts and reviewed full-texts were independently screened by paired researchers (E.N, A.F and G.S) between December 2017 and February 2018 following the inclusion criteria, using EPPI-Reviewer software⁵⁶. Descriptive characteristics of each ontology were extracted, including domain, main areas captured and relevance to human behaviour change. The relevance to human behaviour change was assessed with reference to the overarching scope of the HBCP⁶, addressing the question “What works compared with what, for what behaviours, how well, for how long, with whom, in what setting and why?” and via comparison to terminologies included in taxonomies that have been developed in the domain thus far^{10,11}. The size of each ontology was assessed via the number of classes (i.e. entities), individuals (instances of classes), properties (attributes of a class and relations linking classes) and expressivity (the use of constructs in an underlying logical formalism to computationally encode the ontology) from BioPortal metrics (<https://bioportal.bioontology.org/ontologies>) and by accessing the most recent version (where available, specific version information is documented, otherwise dates of access are used to delineate the version) of the ontology file in the Protégé ontology editor software⁵⁷. Each ontology was checked in Protégé for logical consistency by performing classification with the HermiT reasoner⁵⁸. Counts of textual definitions and counts of full logical definitions were assessed programmatically for each ontology in Python using the Owlready2 library⁵⁹.

Details of the methods used to develop each ontology were extracted from their accompanying papers. This included whether the ontology incorporated existing taxonomies, terminologies (e.g. medical

vocabularies SNOMED⁵⁴ and MedDRA⁵⁵) or ontologies and whether development was informed by user feedback (e.g. expert review of the ontology) and/or by data-driven methods (e.g. using healthcare data). Disagreements over quality assessments were planned to be resolved by discussion with J.H, although no disagreements were found between reviewers.

Quality assessment of identified ontologies

There are many different ways to evaluate the quality of ontologies^{60,61}. Our objective was to identify ontologies that might be re-usable for the domain of behaviour change and evaluate their content rather than their structure, corresponding to the task that Katsumi and Gruninger call “Search”⁶². It is one of the tenets of the OBO Foundry effort that adherence to a common set of principles for ontology development makes ontologies interoperable and thereby facilitates re-use of content²⁴. Informed by this framework, we generated quality assessment criteria by reviewing the OBO Foundry principles of good ontology practice²⁴, which also correspond to some extent to ontology “good design” methodologies such as OntoClean⁶³.

Ontologies were independently assessed (by E.N and A.F) against six criteria drawn from these principles. First, each ontology was assessed for its use of globally unique, unambiguous Uniform Resource Identifiers (URIs) following best practices⁶⁴ for each class (Yes/To Some Extent/No). URIs are strings of characters used to identify an entity, e.g ‘BFO_0000023’ representing the entity ‘role’ in the Basic Formal Ontology (BFO)¹⁵. Standards-compliant URIs are important as they allow unambiguous identification of terms even if labels are updated and facilitate cross-referencing between ontologies²⁴. Secondly, ontologies were assessed for their use of clear definitions for all terms (i.e non-overlapping terms which lack redundancy¹⁵; Yes/To Some Extent/No). Clear definitions are important to allow ontology users and developers to assess the adequacy of a given class for their own uses²⁴. The definitions included in the ontologies identified in this review were primarily expressed in natural language rather than as formal axioms, but where such axioms were present we assessed them. Thirdly, the classification hierarchy of identified ontologies was assessed, with a clear, hierarchical structure preferred to reflect groupings of terms²⁴ (Yes/To Some Extent/No).

Fourthly, we assessed the logical consistency (absence of contradictory axioms⁶⁵) of each ontology by processing the files through the HermiT reasoner⁶⁶ on Protégé software. Fifthly, where possible, we included information about whether the ontology developers conducted an explicit evaluation of their own ontology, and if so what method they used to do this. Finally, ontologies were assessed as to whether they were maintained or not (Yes/No). Maintenance is usually conducted by a community to reflect changes in scientific consensus over time and is essential to ensure the continued relevance of a given ontology²⁴. No disagreements were found between reviewers.

Author contributions

All authors approved the review protocol. EN & AF performed searches. EN, AF, GS performed screening with JH providing feedback. EN & AF performed data extraction and quality assessment, with JH acting as a third reviewer for any conflicts. EN and AF wrote the first draft, with all authors contributing to drafts and approving the final version of the manuscript.

Data availability

The data that support the findings of this study are available from the corresponding author upon request.

References

- 1 Michie, S., Van Stralen, M. M. & West, R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation Science* **6**, 42 (2011).
- 2 The Cochrane Collaboration. *The Cochrane Library*, <<http://www.cochranelibrary.com/>> (2018).
- 3 Michie, S. & Johnston, M. Optimising the value of the evidence generated in implementation science: the use of ontologies to address the challenges. *Implementation Science* **12**, 131 (2017).
- 4 Michie, S., West, R., Campbell, R., Brown, J. & Gainforth, H. *ABC of behaviour change theories*. Silverback Publishing (2014).

- 5 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
Psychology Review* **9**, 323-344 (2015).
- 6 Michie, S. *et al.* The Human Behaviour-Change Project: harnessing the power of artificial
intelligence and machine learning for evidence synthesis and interpretation. *Implementation
Science* **12**, 121 (2017).
- 7 Ioannidis, J. P. *et al.* Increasing value and reducing waste in research design, conduct, and
analysis. *The Lancet* **383**, 166-175 (2014).
- 8 Stavri, Z. & Michie, S. Classification systems in behavioural science: current systems and
lessons from the natural, medical and social sciences. *Health Psychology Review* **6**, 113-140
(2012).
- 9 Hollands, G. J. *et al.* The TIPPMME intervention typology for changing environments to
change behaviour. *Nature Human Behaviour* **1**, 0140 (2017).
- 10 Michie, S. *et al.* The behavior change technique taxonomy (v1) of 93 hierarchically clustered
techniques: building an international consensus for the reporting of behavior change
interventions. *Annals of Behavioral Medicine* **46**, 81-95 (2013).
- 11 Carey, R. N. *et al.* Describing the 'how' of behaviour change: a taxonomy of modes of
delivery. *UK Society for Behavioural Medicine Conference* (2016).
- 12 Michie, S. *et al.* From theory-inspired to theory-based interventions: A protocol for
developing and testing a methodology for linking behaviour change techniques to theoretical
mechanisms of action. *Annals of behavioral medicine* <https://doi.org/10.1007/s12160-016-9816-6> (2016).
- 13 Howlett, N., Trivedi, D., Troop, N. A. & Chater, A. M. Are physical activity interventions for
healthy inactive adults effective in promoting behavior change and maintenance, and which
behavior change techniques are effective? A systematic review and meta-analysis.
Translational Behavioral Medicine, doi:10.1093/tbm/iby010 (2018).
- 14 Michie, S., West, R., Sheals, K. & Godinho, C. A. Evaluating the effectiveness of behavior
change techniques in health-related behavior: a scoping review of methods used.
Translational Behavioral Medicine **8**, 212-224 (2018).
- 15 Arp, R., Smith, B. & Spear, A. D. *Building ontologies with basic formal ontology*. (MIT
Press, 2015).
- 16 Busse, J. *et al.* Actually, What Does “Ontology” Mean? *Journal of Computing and
Information Technology* **23**, 29-41 (2015).
- 17 Blanch, A. *et al.* Ontologies about human behavior: A review of knowledge modeling
systems. *European Psychologist* **22**, 180 (2017).
- 18 Larsen, K. R. *et al.* Behavior change interventions: the potential of ontologies for advancing
science and practice. *Journal of Behavioral Medicine* **40**, 6-22 (2017).
- 19 Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nature Genetics* **25**,
25 (2000).
- 20 Bauer, S. Gene-category analysis, in *The Gene Ontology Handbook*, 175-188 (Springer,
2017).
- 21 Kraker, P. *et al.* The Vienna principles: a vision for scholarly communication in the 21st
century. *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare*
69, 436-446 (2016).
- 22 Noy, N. F. *et al.* BioPortal: ontologies and integrated data resources at the click of a mouse.
Nucleic Acids Research **37**, W170-W173 (2009).
- 23 Smith, B. *et al.* The OBO Foundry: coordinated evolution of ontologies to support biomedical
data integration. *Nature Biotechnology* **25**, 1251 (2007).
- 24 OBO Foundry. Principle: Overview, <<http://www.obofoundry.org/principles/fp-000-summary.html>> (2018).
- 25 Poldrack, R. A. *et al.* The cognitive atlas: toward a knowledge foundation for cognitive
neuroscience. *Frontiers in Neuroinformatics* **5**, 17 (2011).
- 26 Hastings, J., Smith, B., Ceusters, W., Jensen, M. & Mulligan, K. in *ICBO 2012: 3rd
International Conference on Biomedical Ontology*. 1-5 (2012).

- 27 Gkoutos, G. V., Schofield, P. N. & Hoehndorf, R. in *International Review of Neurobiology* Vol. 103 69-87 (Elsevier, 2012).
- 28 Ceusters, W. & Smith, B. Foundations for a realist ontology of mental disease. *Journal of Biomedical Semantics* **1**, 10 (2010).
- 29 Jensen, M. *et al.* The neurological disease ontology. *Journal of Biomedical Semantics* **4**, 42 (2013).
- 30 Schriml, L. M. *et al.* Disease Ontology: a backbone for disease semantic integration. *Nucleic Acids Research* **40**, D940-D946 (2011).
- 31 Schriml, L. M. *et al.* GeMInA, Genomic Metadata for Infectious Agents, a geospatial surveillance pathogen database. *Nucleic Acids Research* **38**, D754-D764 (2009).
- 32 Mattingly, C. J., McKone, T. E., Callahan, M. A., Blake, J. A. & Hubal, E. A. C. Providing the missing link: the exposure science ontology. *Environmental Science & Technology* (ACS Publications, 2012).
- 33 Turner, J. A. & Laird, A. R. The cognitive paradigm ontology: design and application. *Neuroinformatics* **10**, 57-66 (2012).
- 34 Gil, R., Virgili-Gomá, J., García, R. & Mason, C. Emotions ontology for collaborative modelling and learning of emotional responses. *Computers in Human Behavior* **51**, 610-617 (2015).
- 35 Hastings, J., Ceusters, W., Smith, B. & Mulligan, K. Dispositions and processes in the Emotion Ontology. *International Conference on Biomedical Ontology* (2011).
- 36 Pesquita, C., Ferreira, J. D., Couto, F. M. & Silva, M. J. The epidemiology ontology: an ontology for the semantic annotation of epidemiological resources. *Journal of Biomedical Semantics* **5**, 4 (2014).
- 37 Hicks, A., Hanna, J., Welch, D., Brochhausen, M. & Hogan, W. R. The ontology of medically related social entities: recent developments. *Journal of Biomedical Semantics* **7**, 47 (2016).
- 38 Phan, N., Dou, D., Wang, H., Kil, D. & Piniewski, B. Ontology-based deep learning for human behavior prediction with explanations in health social networks. *Information Sciences* **384**, 298-313 (2017).
- 39 Bickmore, T. W., Schulman, D. & Sidner, C. L. A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology. *Journal of Biomedical Informatics* **44**, 183-197 (2011).
- 40 Prochaska, J. O. & Velicer, W. F. The Transtheoretical Model of Health Behavior Change. *American Journal of Health Promotion* **12**, 38-48 (1997).
- 41 Hoffmann, T. C. *et al.* Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* **348**, g1687 (2014).
- 42 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (DSM-5®)*. (American Psychiatric Pub, 2013).
- 43 OBO Foundry. *Basic Formal Ontology*, <<http://www.obofoundry.org/ontology/bfo.html>> (2018).
- 44 Ceusters, W. An information artifact ontology perspective on data collections and associated representational artifacts. *MIE*. 68-72 (2012).
- 45 Courtot, M. *et al.* MIREOT: The minimum information to reference an external ontology term. *Applied Ontology* **6**, 23-33 (2011).
- 46 Richard, M., Aimé, X., Krebs, M.-O. & Charlet, J. Enrich classifications in psychiatry with textual data: an ontology for psychiatry including social concepts. *MIE*. 221-223 (2015).
- 47 Wilkinson *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, **3**. 160018. (2016).
- 48 Shneiderman, B. *Designing the user interface: strategies for effective human-computer interaction*. (Pearson Education India, 2010).
- 49 Arksey, H. & O'Malley, L. Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology* **8**, 19-32 (2005).
- 50 Simera, I. *et al.* Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Medicine* **8**, 24, doi:10.1186/1741-7015-8-24 (2010).

- 51 Matentzoglou, N., Malone, J., Mungall, C. & Stevens, R. MIRO: guidelines for minimum information for the reporting of an ontology. *Journal of Biomedical Semantics* **9**, 6, doi:10.1186/s13326-017-0172-7 (2018).
- 52 Tricco, A.C. *et al.* PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine* [Epub ahead of print]. doi:10.7326/M18-0850
- 53 Norris, E., Finnerty, A. N., Hastings, J., Stokes, G. & Michie, S. Advancing methods to develop behaviour change interventions: a review of relevant ontologies. *PROSPERO International prospective register of systematic reviews*, CRD42017079990 (2017).
- 54 De Silva, T. S., MacDonald, D., Paterson, G., Sikdar, K. C. & Cochrane, B. Systematized nomenclature of medicine clinical terms (SNOMED CT) to represent computed tomography procedures. *Computer Methods and Programs in Biomedicine* **101**, 324-329 (2011).
- 55 Brown, E. G., Wood, L. & Wood, S. The medical dictionary for regulatory activities (MedDRA). *Drug Safety* **20**, 109-117 (1999).
- 56 Thomas, J., Brunton, J. & Graziosi, S. EPPI-Reviewer 4.0: software for research synthesis. EPPI-Centre Software. London: Social Science Research Unit. *Institute of Education, University of London* (2010).
- 57 Noy, N. F. *et al.* Protégé 2000: an open-source ontology-development and knowledge-acquisition environment. in *AMIA Annual Symposium proceedings. AMIA Symposium*. 953-953 (2003).
- 58 Shearer, R., Motik, B. & Horrocks, I. HermiT: A Highly-Efficient OWL Reasoner. *OWLED*, **432**, 91-101 (2008).
- 59 Lamy, J.B. Owlready: Ontology-oriented programming in Python with automatic classification and high-level constructs for biomedical ontologies. *Artificial Intelligence in Medicine*. **80**, 11-28 (2017).
- 60 Vrandečić, D. *Ontology Evaluation* In: Staab, S. & Studer, R. (Eds). Handbook on Ontologies. Springer, Berlin (2009).
- 61 Amith, M., He, Z., Bian, J., Antonio Lossio-Ventura, J. & Tuo, C. Assessing the practice of biomedical ontology evaluation: Gaps and opportunities. *Journal of Biomedical Informatics*. **80**, 1-13. (2018).
- 62 Katsumi, K. & Gruninger, M. Choosing ontologies for reuse. *Applied Ontology*. **12**, 195-221. (2017).
- 63 Guarino, N. & Welty, C. Evaluating ontological decisions with OntoClean. *Communications of the ACM*. **45**, 61-65. (2002).
- 64 McMurry, J.A. *et al.* Identifiers for the 21st century: How to design, provision, and reuse persistent identifiers to maximise utility and impact of life science data. *PLOS Biology*, **15**, 6, e2001414 (2017).
- 65 Horridge, M., Parsia, B. & Sattler, U. Explaining inconsistencies in OWL ontologies. *International Conference on Scalable Uncertainty Management*. (2009).
- 66 Glimm, B., Horrocks, I., Motik, B., Stoilos, G. & Wang, Z. HermiT: An OWL 2 Reasoner. *Journal of Automated Reasoning*. **54**, 245-269. (2014).

Acknowledgements

We thank the Wellcome Trust for funding the project which gave rise to this paper: ‘The Human Behaviour-Change Project: Building the science of behaviour change for complex intervention development’ (201,524/Z/16/Z). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. Thanks to Elise Crayton, Stephanie Stanton-Fay, Holly Walton and Alison Wright for providing comments on an earlier draft.

Competing interests

The authors declare no competing interests.

Box 1. Glossary of terms related to ontologies.

Term	Definition
Annotation	Process of manually or automatically identifying entities from an ontology within documents ¹⁵
Basic Formal Ontology (BFO)	A small, upper level ontology which provides a structure for development by dividing entities into two categories of continuants and occurrents ¹⁵
BioPortal	Repository (portal) for the hosting and maintenance of biomedical-related ontologies ²²
Continuant	Entities within an ontology that continue to exist self-identically over time, e.g. objects and spatial regions ¹⁵
Entity	Objects, attributes and/or processes represented in a given ontology ¹⁵
Occurrent	Entities within an ontology that extend over time, e.g. processes ¹⁵
Ontology	A data structure of: i) unique identifiers representing types of entity, ii) labels and definitions corresponding to these identifiers, and ii) specified relationships between the entities ¹⁵⁻¹⁷
Open Biological and Biomedical Ontology (OBO) Foundry	Repository (portal) for the hosting and maintenance of ‘gold standard’ ontologies ²³ adhering to clearly defined principles of best practice ²⁴
Taxonomy	Classification system representing hierarchical, ordered categories of entities e.g Behaviour Change Techniques Taxonomy (BCTTv1) ^{8,10}
Terminology	Body of terms representing a particular domain e.g. medical vocabularies such as SNOMED ⁵⁴ and MedDRA ⁵⁵
Typology	Classification system representing general types of entities e.g TIPPM typology ⁹

Table 1. Overview of ontologies related to human behaviour change.

Name	Domains	URL	Version accessed	No. of Classes	No. of Individuals	No. of Properties	Expressivity	BioPortal ²²	OBO Foundry ²³	Co for cha
Cognitive Atlas (COGAT) ²⁵	Cognitive neuroscience, mental processes.	https://bioportal.bioontology.org/ontologies/COGAT	15.08.16 (v. 0.3.1)	3639	0	6	ALC	Y	N	N
Cognitive Paradigm Ontology (COGPO) ³³	Cognitive and behavioural experiments	https://bioportal.bioontology.org/ontologies/COGPO	16.12.10	400	25	58	SROIN(D)	Y	N	N
EmotionsOnto ³⁴	Emotions	http://rhizomik.net/html/ontologies/emotions&cognitionontology/	05.2015 (v. 3)	41	0	9	ALN(D)	N	N	N
Emotion Ontology (MFOEM) ³⁵	Emotions	https://bioportal.bioontology.org/ontologies/MFOEM	26.10.17 (v. 27)	902	19	29	SROIQ	Y	Y	N
Epidemiology Ontology (EPO) ³⁶	Epidemiology, statistics	http://www.obofoundry.org/ontology/epo.html	21.06.13 (v. 3)	191	0	4	ALC	N	Y	N
Exposure Ontology (EXO) ³²	Exposure science, genomics, toxicology	http://www.obofoundry.org/ontology/exo.html	10.04.17 (v. 7)	148	0	11	SR	Y	Y	N
Health Behaviour Change Ontology (HBCO) ³⁹	Behaviour change, automated dialogue systems	https://sourceforge.net/p/hbco/code/HEAD/tree/trunk/ontology/	03.06.2011	92	357	82	SROIN(D)	N	N	N
Human Disease Ontology (DOID) ³⁰	Disease	http://www.obofoundry.org/ontology/doid.html	02.03.18 (v. 558)	12498	0	15	SHI	Y	Y	N
Mental Disease Ontology (MFOMD) ²⁸	Mental disease, mental processes	http://www.obofoundry.org/ontology/mfomd.html	21.09.17 (v. 1)	1127	19	48	SROIQ	Y	Y	N
Mental Functioning Ontology (MF) ²⁶	Mental functioning, mental processes	http://www.obofoundry.org/ontology/mf.html	26.10.17 (v. 15)	692	19	23	SROIQ	Y	Y	N
Neuro Behaviour Ontology (NBO) ²⁷	Behavioural processes and phenotypes	http://www.obofoundry.org/ontology/nbo.html	11.05.18	1036	0	15	ALCF(D)	Y	Y	N

Neurological Disease Ontology (ND) ²⁹	Neurological disease, aetiology	https://github.com/addiehl/neurological-disease-ontology/blob/master/src/ontology/ND.owl	06.05.16 (v. 4.9)	671	19	61	SROIF	N	N	N
Ontology of Medically Related Social Entities (OMRSE) ³⁷	Social entities related to healthcare, demographics, social roles	http://www.obofoundry.org/ontology/omrse.html	16.07.18 (v. 37)	369	33	76	SROIQ(D)	Y	Y	N
Semantic Mining of Activity, Social and Health Data (SMASH) ³⁸	Healthcare data, sustained weight loss	https://bioportal.bioontology.org/ontologies/SMASH	02.06.15 (v. 6)	189	87	144	SHOIQ(D)	Y	N	N
Symptom Ontology (SYMP) ³¹	Symptom, disease	http://www.obofoundry.org/ontology/symp.html	11.05.18 (v. 15)	942	0	1	AL	Y	Y	N

Notes: Information correct as of data extraction in February 2018. Information on classes, properties and individuals taken from BioPortal metrics where available and most recent version ontology file where not available.

Expressivity meanings: Each letter in the expressivity code corresponds to different aspects of ontology logic. ALC – This is the base description logic language, “Attributive Language with Complements”. It includes class constructors for negation, union, and intersection, and existential and universal qualifiers for role restrictions; AL – The base description logic language, without negation; ALN(D) – The base description logic language with number restrictions (N) and data properties (D) – which map to e.g. dates, numbers etc. and allow simple numeric computations; ALCF(D) – The base description logic language with specification of functional roles (F), i.e. properties that can take only one filler, and data properties (D); SR – The base description logic language with transitive roles, role hierarchy, disjoint roles, and some additional allowed complexities in the definitions of roles; SHI – The base description logic language with transitive roles, inverse roles, and role hierarchy; SROIQ – As per SR, as well as inverse roles (I), qualified number restrictions (Q) and nominals (O); SROIQ(D) – As per SROIQ, as well as data properties; SROIF – As per SR, as well as nominals (O), inverse roles (I) and functional roles (F); SROIN(D) – As per SR, as well as nominals (O), inverse roles (I), unqualified number restrictions (N), and data properties (D). Code explanations extracted from <http://www.cs.man.ac.uk/~ezolin/dl/>.

Table 2. Methods used and quality assessment of identified ontologies.

Name	Ontology Methods					Quality assessment					
	Uses existing taxonomies	Uses existing terminologies	Uses existing ontologies	User feedback	Data-driven	Unique URIs	Clear definitions	Clear structure	Logically consistent	Evaluated	Maintained
Cognitive Atlas (COGAT) ²⁵	N	N	Y	Y	N	Y	To some extent	To some extent	Y	N	Y
Cognitive Paradigm Ontology (COGPO) ³³	Y	N	Y	Y	N	Y	To some extent	Y	Y	N	N
EmotionsOnto ³⁴	N	N	Y	N	Y	N	N	To some extent	Y	Y	N
Emotion Ontology (MFOEM) ³⁵	N	N	Y	N	N	Y	Y	Y	Y	N	Y
Epidemiology Ontology (EPO) ³⁶	N	Y	Y	N	N	Y	N	Y	Y	N	Y
Exposure Ontology (EXO) ³²	N	Y	Y	Y	Y	Y	Y	Y	Y	N	Y
Health Behaviour Change Ontology (HBCO) ³⁹	To some extent	N	Y	Y	N	N	N	To some extent	N	N	N
Human Disease Ontology (DOID) ³⁰	N	Y	Y	Y	N	Y	Y	Y	Y	N	Y
Mental Disease Ontology (MFOMD) ²⁸	N	Y	Y	N	N	Y	Y	Y	Y	N	Y
Mental Functioning Ontology (MF) ²⁶	N	N	Y	N	N	Y	Y	Y	Y	N	Y
Neuro Behaviour Ontology (NBO) ²⁷	N	N	Y	N	Y	Y	Y	Y	Y	N	Y
Neurological Disease Ontology (ND) ²⁹	N	Y	Y	N	N	To some extent	Y	Y	Y	N	Y

Ontology of Medically Related Social Entities (OMRSE) ³⁷	Y	N	Y	Y	N	Y	Y	Y	Y	N	Y
Semantic Mining of Activity, Social and Health Data (SMASH) ³⁸	N	N	N	N	Y	N	N	Y	N	Y	N
Symptom Ontology (SYMP) ³¹	N	Y	Y	N	N	Y	Y	Y	Y	N	Y

Notes: Information taken from ontology files and published reports for each ontology

Figure 1. Flow chart for identification of ontologies.

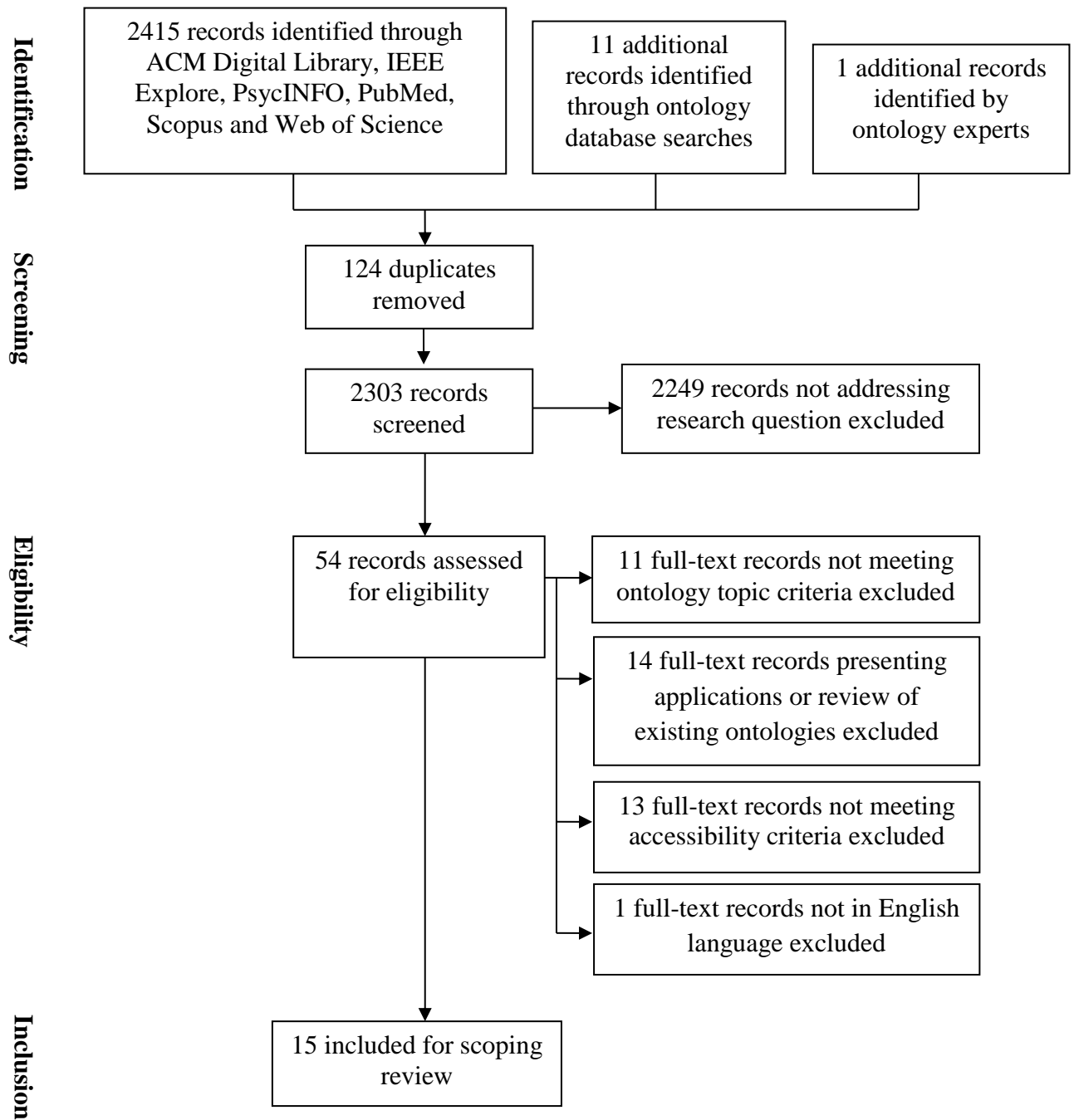


Figure 2. Excerpt from Health Behaviour Change Ontology (HBCO)³⁹

