Article:
Hollands, GJ, Bignardi, G, Johnston, M et al. (7 more authors) (2017) The TIPPME intervention typology for changing environments to change behaviour. Nature Human Behaviour, 1. 0140. ISSN 2397-3374

https://doi.org/10.1038/s41562-017-0140

TITLE PAGE

TITLE

The TIPPME intervention typology for changing environments to change behaviour

AUTHORS

***Gareth J Hollands, PhD[1]; Giacomo Bignardi, MRes[1]; Marie Johnston, PhD[2]; Michael P Kelly, PhD[3]; David Ogilvie, PhD[4]; Mark Petticrew, PhD[5]; Andrew Prestwich, PhD[6]; Ian Shemilt, BA[7]; Stephen Sutton, PhD[3]; Theresa M Marteau, PhD[1].

*CORRESPONDING AUTHOR **(gareth.hollands@medschl.cam.ac.uk)**

AFFILIATIONS. ADDRESSES

(1) Behaviour and Health Research Unit, University of Cambridge.
Institute of Public Health, Robinson Way, Cambridge CB2 0SR, UK.
(2) Institute of Applied Health Sciences, University of Aberdeen.
Polwarth Building, Aberdeen AB25 2ZD, UK.
(3) Department of Public Health and Primary Care, University of Cambridge.
Institute of Public Health, Robinson Way, Cambridge CB2 0SR, UK.
(4) MRC Epidemiology Unit, University of Cambridge.
University of Cambridge School of Clinical Medicine, Box 285 Institute of Metabolic Science
Cambridge Biomedical Campus, Cambridge CB2 0QQ, UK.
(5) Department of Social and Environmental Health Research, London School of Hygiene &
Tropical Medicine.
Keppel Street, London WC1E 7HT, UK.
(6) School of Psychology, University of Leeds.
University of Leeds, Leeds LS2 9JT, UK.
(7) EPPI-Centre, UCL Institute of Education.
University College London, 18 Woburn Square, London WC1H 0NR, UK.

34 **The TIPPME intervention typology for changing environments to change behaviour**

35

36 **ABSTRACT**

37

38 Reflecting widespread interest in concepts of 'nudging' and 'choice architecture', there is
39 increasing research and policy attention on altering aspects of the small-scale physical environment,
40 such as portion sizes or product positioning, to change health-related behaviour at population-level.
41 There is, however, a lack of clarity in characterising these interventions, and no reliable framework
42 incorporating standardised definitions. This hampers both the synthesis of cumulative evidence
43 about intervention effects, and the identification of intervention opportunities. To address this, a
44 new tool, TIPPME (Typology of Interventions in Proximal Physical Micro-Environments) has been
45 developed, here applied to the selection, purchase and consumption of food, alcohol and tobacco.
46 This provides a framework to reliably classify and describe, and enable more systematic design,
47 reporting and analysis of, an important class of interventions. In doing so, it makes a distinct
48 contribution to collective efforts to build the cumulative evidence-base for effective ways of
49 changing behaviour across populations.

50

51

52

53

54

55

56

57

58

Unhealthy patterns of food, alcohol and tobacco consumption are major contributors to the burden of non-communicable diseases – currently accounting for more than two thirds of deaths worldwide[1,2]. It is now widely recognised that the physical environments that surround us exert considerable influence on these patterns of consumption, and that changing these environments holds corollary potential as a catalyst for changing consumption. Whilst not new, the idea that behaviour can be changed in predictable ways, by changing the environments within which people make choices – 'choice architecture'[3] - has gained traction globally among the public, the research community, and policymakers[4,5]. However, despite the recent popularisation and intuitive appeal of these approaches, there has been an absence of definitional and conceptual clarity in characterising such interventions, particularly regarding applications to public health. The absence of a reliable framework that incorporates standardised labels and definitions has hampered the synthesis of cumulative evidence about intervention effects, resulting in an evidence-base that remains uneven and uncertain. It has also hindered the identification and discussion of opportunities to intervene to change environments.

In response to these observations, we present and provide guidance for a new tool – TIPPME (Typology of Interventions in Proximal Physical Micro-Environments) - that aims to improve researchers' and practitioners' ability to clearly and consistently classify and describe an important class of behaviour change interventions related to concepts of 'nudging' and 'choice architecture'. The focus of the typology is on interventions that involve altering aspects of physical micro-environments to change health-related behaviour, here specifically applied to the selection, purchase and consumption of food, alcohol and tobacco products. A more detailed discussion of definitions and concepts follows below, but in essence, these interventions involve changing characteristics of products themselves and the environment in which they are available, within places such as shops, restaurants, bars, and workplaces. Examples include altering the portion size of food, alcohol and tobacco products, and changing their availability or positioning within an

environment, such as providing additional healthier options to select from or placing less healthy options further away from potential consumers.

Aims of TIPPME

To provide a framework for reliably classifying and describing ways in which interventions can alter proximal physical micro-environments to change selection, purchase and consumption of food, alcohol and tobacco products, in order to:

i) Facilitate the synthesis of cumulative evidence about the effects of interventions that can be mapped on to the areas of influence or responsibility of different potential actors (e.g. industry, policy makers, public), including supporting clearer reporting of intervention content in primary and secondary research; and,

ii) Facilitate identification and discussion of a broader range of opportunities for interventions to be developed, implemented and evaluated. This is potentially useful for researchers and those in positions to directly alter, or advocate for changes to, commercial, public sector or domestic environments.

Focus of TIPPME

We define the focus of this typology as:

Interventions or ways to alter the properties or the placement of objects or stimuli within proximal (sensorily perceptible) physical micro-environments, to elicit particular behaviours among people within those environments. These interventions are implemented within the same environment as that in which the target behaviour is performed, and are not designed to be interactive or tailored to specific individuals.

Our choice of terminology regarding 'proximal physical micro-environments' is intended to reflect the spatial focus of this class of interventions relative to the people exposed to them. It draws upon a conceptual distinction made within the ANGELO (Analysis Grid for Environments Linked to Obesity) framework[6] between two levels of environment, micro- and macro-. Micro-environments are settings which people use for specific purposes (e.g. shops, restaurants and bars) and where they interact directly with objects and stimuli within those environments. In contrast, macro-environments are the higher-level systems and infrastructure that influence the characteristics of micro-environments and the relationship between them (e.g. the availability of micro-environments themselves, such as the geographical distribution of shops, restaurants and bars in a given area).

ANGELO additionally distinguishes between four types of environment: physical, economic, political and socio-cultural. As we are concerned with the consumption of food, alcohol and tobacco products that are themselves objects within it, the stated focus of TIPPME is on the physical micro-environment. We have not attempted concurrently to map economic, political and socio-cultural environments, though we acknowledge their importance and the complex interactive relationships between them, and between interventions and outcomes. For example, these other environments may be manifest in any changes made to physical environments (e.g. political environments may influence physical environments) or changes made to physical environments may impact upon them (e.g. providing information may influence social norms). As well as directly influencing the nature of physical environments, economic, political and socio-cultural environments also determine the background conditions in which people are exposed to physical environments, such as times when the physical environments can be accessed, and the economic costs that are imposed upon them. Applying a sociological lens, social structures constrain and enable the actions of individual human agents and set the limits of behavioural possibilities. Humans operate in an environment which is simultaneously social, biological, and physical[7]. Here our focus is on the physical, while acknowledging this wider set of parameters.

Because physical micro-environments can be very large and encompass a wide range of functions and purposes for the people in them (for example, neighbourhoods or streets), this term is not specific enough to capture the focus of the interventions we aim to characterise. The addition of 'proximal' reflects our conceptual focus, as these interventions are typically implemented close (spatially and temporally) to the point of decision or performance of the people exposed to them, in order to influence behaviour enacted in that same physical environment. We have bounded the parameters of the physical environments characterised by our typology to those that are sensorily perceptible (i.e. able to be seen, heard, smelt, touched, or tasted) by intervention recipients. In combination with the other elements of our definition, this is intended to give an approximate indication of the likely scale of the interventions of interest, given that precisely and accurately quantifying the range of distances is not practicable. Finally, the focus of this typology excludes interventions that are designed to be interactive or tailored, meaning those in which the intervention content is not standardised for all recipients and is intended or enabled to vary dependent on their characteristics or responses. This may result from an interaction with a person or machine, such as a cafeteria worker or a computer-based system providing personalised nutritional guidance based on food purchasing patterns, demographic characteristics or responses to questions. While interventions included in TIPPME are not necessarily non-interactive - in so far as people could in theory interact with them and change their content - they are not designed to elicit such interaction.

The importance of interventions in proximal physical micro-environments

Interventions in proximal physical micro-environments have significant potential to change behaviour to improve population health[8]. This is reflected in current policy and research interest. These interventions have key advantages over many other types of behaviour change interventions. First, the nature of altering characteristics of physical environments means that these interventions have the potential to shape the behaviour of all those exposed to that environment without the need

for interpersonal interaction. This means that once an intervention has been developed and implemented, there are likely minimal ongoing resource costs associated with its continued use. Second, because physical environments have the potential to be modified in a consistent and directly measurable way, an intervention can be readily and reliably transferred to other locations, and scaled up in its application to reach larger populations. Third, because these interventions typically involve altering cues located proximally in time and space to the behaviour, their effects are likely less reliant on people purposefully, consciously engaging with the intervention over time[9], or on high levels of personal agency[10]. This means that they may be less affected by differential, often socially patterned, cognitive or motivational resources. They therefore, in theory, have the potential to be effective across the populations to which they are applied, without widening existing health inequalities. Such potential is reflected in evidence suggesting that interventions that alter the environments to which people are exposed may be less likely to widen inequalities than individual-level education and counselling[11]. Finally, there is emerging evidence that this kind of public health intervention is more acceptable to the public than economic interventions such as taxes on products[12,13], public acceptability being a key determinant of whether an intervention is implemented[14].

Previous attempts to characterise the proximal physical micro-environment

In addition to the ANGELO framework[6], there have been several complementary research efforts that make reference to small-scale physical environments. The Behaviour Change Technique (BCT) Taxonomy[15] aims at comprehensively describing behaviour change techniques, including 'restructuring the physical environment', although it does not further classify interventions within this category. The Intervention Mapping approach[16] describes a series of steps for developing interventions, and includes 'nudging' as one possible approach to changing determinants underlying behaviour, but does not disassemble this concept in terms of specific intervention content. Other

work has focused on classifying characteristics of 'nudging' or 'choice architecture' interventions (e.g.[17,18]), but these typically concern broad theoretical principles and do not describe ways of changing physical environments in any detail. Previous attempts to map features of the physical environment that cue our behaviour or ways in which it can be changed[19-22] are unable to address our aims adequately, as they are insufficiently detailed, not systematically developed and assessed, or are not applied to consumptive health-related behaviours.

In an earlier phase of this research, we developed a provisional typology that focused specifically on the ways in which small-scale physical environments have been altered to influence food, alcohol, tobacco and physical activity behaviours[23,24]. This was derived from a large-scale systematic scoping review of the research literature on 'choice architecture' interventions, intended to map the parameters of previous empirical research and provide a conceptual map of the evidence base, in order to delineate and characterise more specific intervention types. In the current paper we describe further development of this work, introducing TIPPME (Typology of Interventions in Proximal Physical Micro-Environments), which is intended to improve upon and replace the provisional typology. Such development was needed because the nature of the provisional typology was determined by the extant research literature, and was therefore not designed to be applied beyond organising that specific body of literature. Additionally, the early stages of developing TIPPME (see Methods, Stages 1-3) identified various conceptual issues with the provisional typology that undermined its validity and usefulness. TIPPME therefore represents an attempt to produce a more generalisable typology with a conceptually and theoretically coherent structure that can accommodate both interventions that have been developed and tested, and those that exist only in theory. Using the original scoping review process as a platform, it has been shaped in accordance with the collective understanding of experienced researchers and practitioners, with the aim of producing as complete and parsimonious account of the phenomena of interest as possible.

Whilst related to these concepts, the focus of TIPPME has been deliberately distanced from the terminology of 'nudging' and 'choice architecture', this being potentially contentious in terms of how it has been bound to particular political and philosophical positions, and which has been inconsistently interpreted and applied. As Oliver (2015) highlights[25], for a nudge to align with the founding principles of libertarian paternalism[3], it should fulfil a set of essential criteria (e.g. that it is not regulatory, and does not rely on rational reasoning processes). Because interventions that are claimed to represent nudges often do not meet these criteria, continued imprecise usage of the term has resulted in the concept it denotes being obfuscated and confusion around its meaning and potential policy value[26]. While interventions within TIPPME may map on to the concept of nudging in some respects, this is not a necessary feature of the typology. It is therefore instead linked to the more generalised and readily definable concept of the physical environment and the ways in which this can be altered to change behaviour.

RESULTS

This section describes the final version of TIPPME that resulted from the seven sequential stages of its development. Specific results from each stage of that process, including the results of reliability testing exercises, are described in the Methods section.

The final, complete version of TIPPME (Typology of Interventions in Proximal Physical Micro-Environments) is provided in Supplementary Information. It is also available at http://www.bhru.iph.cam.ac.uk/resources/TIPPME (with training materials also provided at this website) and at https://doi.org/10.6084/m9.figshare.5053672. This complete version includes full definitions and guidance, a set of instructions for use, and provides examples of interventions within

each category in the typology. For illustration only, a simplified version of TIPPME is presented in Figure 1. TIPPME includes and encompasses interventions that meet our stated definition of proximal physical micro-environment interventions. In terms of the wording used and examples provided, we have here applied it specifically to the selection, purchase and consumption of food, alcohol and tobacco, acknowledging that it could potentially be applied and adapted to other behaviours (see Discussion).

PLEASE INSERT FIGURE 1 HERE

Figure 1. Simplified version of TIPPME (Typology of Interventions in Proximal Physical Micro-Environments), for changing selection, purchase and consumption of food, alcohol and tobacco (see Supplementary information for full version)

TIPPME comprises a matrix classification structure defining six intervention types and three different spatial foci. The rows of the typology represent different intervention types, i.e. ways in which the proximal physical micro-environment can be altered to elicit changes in behaviour. There are six different intervention types (rows), namely: Availability; Position; Functionality; Presentation; Size; Information. These six intervention types can be aggregated into two higher-order classes of intervention: i) those that involve altering the placement of objects or stimuli within proximal physical micro-environments, and ii) those that involve altering the properties of objects or stimuli within proximal physical micro-environments, indicated by the column on the left edge of the figure. The typology also distinguishes between three intervention foci representing differences in the spatial focus of interventions: Product; Related objects; Wider environment. The combination of rows and columns means there are 18 possible intervention categories that can be applied to describe an intervention.

DISCUSSION

The Typology of Interventions in Proximal Physical Micro-Environments, TIPPME, provides a means of reliably classifying and describing an important class of interventions to change health-related behaviour across populations. TIPPME has benefited from an extensive, iterative and explicit development process that included reliability testing using a sample of people involved in researching and implementing interventions, representing those ultimately likely to use it. In line with other conceptual frameworks, this framework will, with use, be found to be imperfect, but it represents an agreement that a point of development has been reached where we judge the typology can usefully fulfil its stated aims.

First, it provides a reliable framework for the synthesis of cumulative evidence about the effects of interventions, with the potential for a shared language. Second, it can facilitate systematic thinking about and identification and discussion of a broader range of opportunities for interventions to be developed, implemented and evaluated. In turn, resulting findings can be integrated with a growing cumulative evidence base to facilitate the development of more effective interventions. In more practical terms, this typology can feasibly be used in tasks such as classifying or organising bodies of literature; identifying, framing and bounding primary research, as well as systematic and conceptual reviews; and providing a way of listing possible intervention strategies. It is potentially useful both for researchers and for framing the actions of those in positions to alter or influence commercial, public sector or domestic environments. This could include public health practitioners and policymakers, as well as those advocating for such changes. Finally, while the predominant focus of TIPPME is on ways of altering environments, it may also be informative in attempts to describe physical features of environments that (as opposed to being implemented as interventions) already exist and may influence behaviour accordingly (or may moderate the effectiveness of interventions that are introduced).

TIPPME also contributes to and complements ongoing efforts by the wider research community to build the foundations of a cumulative evidence base by developing domain ontologies to encode and curate research knowledge about the effects of interventions, and enable its more efficient identification, synthesis and use. Such domain ontologies include representation of the common and distinct features (or attributes) of different types of interventions, and of the proposed 'active ingredient(s)' that determine their effectiveness[27-29]. This encompasses representation of the content of interventions[15] and of the mode, or form, of their delivery[30]. TIPPME contributes to these efforts by specifying the common and distinct features of a specific class of interventions. In particular, it delineates these interventions in terms of: (i) their content - this being the proposed 'active ingredient(s)' that elicit the behavioural response - which in this case concerns the alteration of attributes of objects or stimuli within the proximal physical micro-environment, such as their size or position; and (ii) the focus of that content. Ontological relationships both within TIPPME, and between TIPPME and other relevant typologies or taxonomies, are likely to be complex - particularly as there may be variation in the level of explanation or granularity applied in each case. Further development work is therefore needed to clarify and specify the form and structure of these relationships, as well as the ways in which different frameworks may be usefully applied in combination. Relatedly, TIPPME does not at present attempt to delineate the mechanisms of action that underlie each intervention type but, if it fulfils its stated aims, this should facilitate primary and secondary research directed towards furthering understanding of such mechanisms.

Whether TIPPME is viewed as a typology, or, with additional development and validation, a more definitive taxonomy, depends on the epistemological position that one adopts. The way that some authors describe ontologies is unequivocally realist, whereby a properly developed ontology will describe the real world as it is, rather than just as it appears to be to the observer[31]. Contrary to this, phenomenologists such as Schutz[32] view ontologies as theories about the nature of being in the world, and typologies and typifications as the means of seeing and interpreting that world; the

plastic nature of such conceptual constructs is emphasised. A true ontology would take full account of both realist and phenomenalist perspectives and indeed our approach to developing TIPPME has drawn on both of these traditions. In the initial phase of development, we derived provisional types from empirical studies; while in latter phases we have sought to use a mix of realist and other forms of knowledge to refine the typology to align with various priors (including theoretical understandings) about relations between concepts and ideas, either as we imagine the world appears to be, or as we imagine the world should be if it conformed to our prior beliefs.

We judge the level of granularity of TIPPME to be appropriate to fulfil our aims; being relatively simple but enabling discrimination between multiple intervention types. Evidence of how our provisional typology[23] has been used to, for example, frame funding calls, inform policy documents and to characterise interventions in systematic reviews (see Methods), suggests that TIPPME has the potential to be similarly useable. However, its granularity could be increased, should important intervention sub-types or additional characteristics be highlighted. This could be through conducting systematic reviews of specific intervention types in which key intervention characteristics are identified (e.g.[33-35]), or mapping relationships between TIPPME and other classification systems.

TIPPME and the methods used to generate it have several limitations. Our first reliability testing exercise, while demonstrating that TIPPME can be reliably used by those outside of the research team, used a relatively small sample of experts, predominantly academic researchers. Whilst such a sample is likely to be broadly representative of some of the most likely users of the typology, there will be other groups that were inadequately represented. Furthermore, the majority of the development process was conducted by the core research team. While a wide range of disciplinary backgrounds was represented, and many of the group had extensive policy and guidance development experience, the team was weighted towards research expertise. A more rigorous and

comprehensive series of assessments, likely also integrating responses from a wider cross-section of potential user groups, will be required for greater confidence in TIPPME's reliability, particularly for use outside of the research community. A further limitation is that TIPPME is currently only applied to three consumption behaviours, which, while highly important – with metabolic and dietary risk factors linked to food consumption, as well as smoking and alcohol use, all being amongst the most significant risk factors contributing to global disease burden[2] – do not encompass all of the human behaviours that significantly impact on health. Most notably, although it was included in the provisional typology[23], physical activity was not included here. We judged that it was not practicable to include due to it being conceptually distinct. This is because, unlike selection, purchase and consumption of food, alcohol and tobacco, physical activity does not necessarily relate to products that are separable from and placed within a given environment (see Methods, Stage 3 for further details).

While the current behavioural focus of TIPPME limits its generalisability, the typology is intended to be broad in scope so that it could potentially be adapted to apply to other behaviours. When considering the current and potential future scope of TIPPME in terms of the behavioural domains to which it applies, it may be helpful to map its categories to a systematically-developed framework of behaviours, such as Nudelman and Shiloh's taxonomy of health behaviours[36]. Within this taxonomy, TIPPME is currently aligned to nutrition and risk avoidance behaviours, these being related to consumption of products that are linked to non-communicable disease. In theory, TIPPME may be applicable to a wide range of other behaviours, including those unrelated to product consumption (e.g. physical activity or gambling behaviours), those linked to the prevention of communicable disease (e.g. hygiene-related behaviours) and pro-environmental behaviours to mitigate climate change (e.g. energy use or recycling behaviours). Such translation will require specific programmes of development and testing.

In the process of developing TIPPME, some challenging conceptual issues were encountered. Most notably, to ensure TIPPME had the potential to discriminate, it was intended that each intervention type (typology row) would represent a distinct way in which the proximal physical micro-environment can be altered, with any single discrete intervention component being assignable to a single intervention type. However, it was quickly apparent that there was an inevitable degree of conceptual overlap or co-dependence between the different intervention categories. For example, if we consider an intervention in which restaurant patrons are provided with smaller (versus larger) spoons to attempt to reduce dessert consumption, we would expect most users would characterise this as a 'Size' intervention. Manipulating the size of a spoon will also likely affect the way the spoon looks and feels i.e. 'Presentation', however, and so either or both types could feasibly be applied for all 'Size' interventions. While this is a fundamental conceptual issue linked to the nature of the sensory, spatial and morphological attributes of objects and stimuli, for the purposes of the typology it can be addressed by distinguishing between the primary target of the intervention and secondary consequences. As explained in the typology and its accompanying instructions (Supplementary information), assuming a single discrete intervention component or manipulation is present, the user applying TIPPME will aim to identify a primary intervention type that best captures it. Should there be multiple discrete, separable intervention components implemented within the same environment, multiple different intervention types can correspondingly be applied. In practice, reports of interventions will often support judgements of what the primary target of the intervention is via their stated aims and hypotheses, and the way in which they describe intervention content. Importantly, results of the reliability testing exercises suggest that despite these potential challenges, the typology can be consistently applied by users.

Because reliable application of TIPPME, as with any classification system, is dependent on making informed judgements from the information that is provided, the increased attention that is being given to improving the reporting of intervention studies[37,38] will be beneficial. Additionally, we

would hope that TIPPME will enable those reporting on proximal physical micro-environment interventions to specify the primary target of their intervention, thereby reducing future ambiguities. The Typology of Interventions in Proximal Physical Micro-Environments, TIPPME, provides a framework to reliably classify and describe an important class of interventions, and enable more systematic design, reporting and analysis of interventions to change health-related behaviour at population level. In doing so, we propose TIPPME makes a distinct contribution to collective efforts to build the cumulative evidence base for effective ways of changing behaviour across populations.

METHODS

The development of TIPPME is summarised in Table 1 and involved three phases of work: a) identifying the need for a typology; b) developing and elaborating on this typology; and c) reliability testing and finalising. These were completed in seven main stages. Throughout, the development process was integrated with formal and informal discussions between the core research team as well as wider academic networks. The core research team (the authors) comprised ten members with a range of disciplinary backgrounds across public health, health policy, psychology and behavioural science, sociology, and evidence synthesis, and many of the group had extensive policy and guidance development experience. They represented varied expertise in developing and applying prominent typologies or classification systems relating to behaviour change interventions and theory, developing, implementing and evaluating public health and behaviour change interventions in a range of behavioural and population contexts, and developing practice and research reporting guidelines.

428 **Table 1. Development process for TIPPME**

429

Phase	Stage	Methods	Results and actions
a) Identifying need	1. Developing a provisional typology of physical micro-environment interventions	Large-scale systematic scoping review to map available empirical evidence.	Produced provisional typology to configure extant literature. Agreed that further development needed to apply more widely.
	2. Receiving feedback from expert workshops	Two workshops, attended by researchers and practitioners (n=45), involving a questionnaire and group discussion.	Support obtained from attendees for value of further development. Considering other indicators of likely value, research team proceeded with development.
b) Developing and elaborating	3. Generating a preliminary version of TIPPME	Two-day residential meeting of research team, with series of structured discussions.	Generated a preliminary version of TIPPME.
	4. Identifying conceptual and practical problems with a preliminary version	Research team completed intervention description coding task, followed by structured discussion via teleconference.	Produced revised version of TIPPME to be subject to reliability testing.
c) Reliability testing and finalising	5. Reliability testing exercise (i): Coding of intervention descriptions by external experts	External experts (n=33) with backgrounds in public health and behaviour change completed exercise involving coding content of 40 short intervention descriptions.	Demonstrated strong reliability in applying TIPPME using short intervention descriptions. Two-day residential meeting of research team held to discuss findings.
	6. Reliability testing exercise (ii): Coding of intervention descriptions using full-text papers	Four members of the research team completed exercise involving coding content of 24 full-text papers.	Demonstrated strong reliability in applying TIPPME to the coding of full-text papers.
	7. Agreement on a final version of TIPPME	Research team members completed final check of the typology and wording, to ensure clear and consistent throughout. Teleconference held to agree on final version.	Produced final version of TIPPME. See Figure 1 for simplified version and Supplementary information for full version.

430

431

432

433

434

435 **Stage 1. Developing a provisional typology of physical micro-environment interventions**

436

437 **Aim:** To generate a refined definition and provisional typology of choice architecture interventions

438 in physical micro-environments, and to map the available empirical evidence for the effects of these

439 interventions on diet, physical activity, alcohol and tobacco use.

440

441 **Methods:** We conducted a large-scale systematic scoping review, published in 2013[23] In brief, the
442 methods used to develop the provisional typology involved highly sensitive searches of 15
443 electronic literature databases, combined with parallel snowball searches, retrieving over 800,000
444 unique title and abstract records. We used text mining methods to prioritise these records for
445 screening[39] and manually screened over 54,000 prioritised records to identify 346 eligible full-text
446 articles.
447
448 **Results:** Data extracted from these 346 articles (reporting primary evaluation studies and reviews of
449 such studies) were then used to configure, describe and synthesise the key characteristics of
450 interventions. This was an iterative process, incorporating regular discussion among members of the
451 review team, and resulted in the provisional typology in Figure 2, comprising nine types of
452 interventions: Ambience, Functional design, Labelling, Presentation, Sizing, Availability,
453 Proximity, Priming, Prompting.
454
455
456
457 **PLEASE INSERT FIGURE 2 HERE**
458
459 **Figure 2. Typology at end of Stage 1**
460
461
462
463 **Subsequent actions:** The research team sought feedback on the typology from a wider group of
464 potential users concerning its usefulness and the scope for further development, including
465 discussion of conceptual issues that had been identified by the research team over the course of
466 conducting the scoping review.
467
468 **Stage 2. Receiving feedback from expert workshops**
469

Aim: To elicit feedback about issues with understanding and using the provisional typology of choice architecture interventions.

Methods: Two typology development workshops were conducted, attended by a total of 45 participants working in areas of behaviour change and public health intervention. Participants were predominantly in research roles (80%, with 20% in policy or practitioner roles), with a range of disciplinary backgrounds represented (psychology or behavioural science (38%); public health or medicine (20%); nutrition (7%), policy (11%) sociology (2%), other or missing (20%)). To encourage engagement with the provisional typology and elicit feedback, the workshops involved completing the same brief intervention description coding task (classifying 14 intervention descriptions by reference to the provisional typology) followed by a questionnaire assessing perceived value of the work and its development, concluding with a structured group discussion.

Results: There was strong support expressed for the value of developing the provisional typology from academics and practitioners who attended: 95% (41/43 responses) of participants agreed with the statement "developing this typology is valuable and important". There was recurrent feedback that further development would be necessary in order for the typology to be more widely applied. Issues highlighted that would need to be considered in future included: identifying some inconsistencies in the intervention types concerning whether they related to the content or the mechanism of the intervention (for example, one of the intervention types, 'priming', related primarily to a mechanism of effect and was therefore not equivalent to other intervention types); whether the 'labelling' and 'prompting' intervention types were clearly distinct; and the difficulty of coding physical activity interventions. The intervention description coding task was principally intended only to encourage engagement and general feedback within the workshops. However, in line with qualitative feedback received from participants, its results did suggest that typology

categories could be applied consistently (Fleiss' kappa =.83), supporting using the basic structure and content of the provisional typology as a foundation for future development.

Subsequent actions: In preparation for future development of the typology, the initial research team was extended to include two behaviour change experts with expertise in developing and applying prominent typologies of behaviour change interventions and theory, and a public health expert with expertise in environmental determinants of health and the evaluation of public health interventions. It was agreed by the research team that development of a new typology would be valuable, informed by both the workshops and other external indicators of potential value. For example, the provisional typology had been used to frame two calls from a national research funding body (National Institute for Health Research (UK)), had directly informed policy documents and guidelines (e.g. NICE (2014), Department of Health (2015)) and was being used for characterising and defining interventions in several systematic reviews (e.g.[33-35,40,41]).

Stage 3. Generating a preliminary version of TIPPME

Aim: To generate by consensus a preliminary version of TIPPME that would be subject to further development.

Methods: A two-day residential meeting of the core research team (the authors) was held. This meeting comprised a series of structured discussions, informed by feedback received from the expert workshops (Stage 2) and from colleagues within wider networks.

Results: A preliminary version of TIPPME was generated. Principal developments agreed at this stage, representing changes to the provisional typology (Figure 2), were as follows:

i. A new matrix classification structure was created to enable representation of both different intervention types (the rows of the typology), as well as differences in the spatial focus of

the intervention (the columns of the typology). This structural change stemmed from agreement that a more conceptually coherent position would be to view any given physical micro-environment as a set of objects or stimuli that could feasibly all be manipulated. This would also allow greater flexibility in thinking about the range of possible interventions within this space, whether these are only theoretically possible or are represented in the current body of empirical literature.

ii. Intervention type 'Ambience' was removed and its place taken by the 'Presentation' intervention type applied on the scale of the wider environment. This was because the new typology structure means that objects and stimuli within the wider environment are considered subject to the same intervention types as are the products themselves.

iii. Intervention types 'Labelling' and 'Prompting' were subsumed within a generic intervention type pertaining to the communication of explicit textual, numeric or pictorial information. This more inclusive category of information-based interventions – initially named 'Words, Numbers and Pictures' and ultimately 'Information' - was considered more coherent, as previous conceptual distinctions between 'Labelling' and 'Prompting' interventions were unclear. This still allowed differentiation from other intervention types that focus on the alteration of sensory, spatial and morphological characteristics.

iv. Intervention type 'Priming' was removed as there was agreement that this represented a specific mechanism rather than an equivalent intervention type.

v. Notably, physical activity was excluded as a behaviour of interest. It was agreed that it was not practicable to include this in a coherent and concise typology, given that, unlike food, alcohol and tobacco, it does not involve the selection, purchase and consumption of products that are separable from and placed within a given environment. It would be possible to adapt the typology to physical activity, with the equivalent of the target product or object being the physical space in which, or on which, the physical activity is performed. This physical space may be a permanent part of, or the whole of, the proximal physical micro-environment

547 itself. However, adapting the typology to physical activity would be complex and require its
548 own specific explanation and translation.
549
550 **Stage 4. Identifying conceptual and practical problems with a preliminary version**
551
552 **Aim:** To attempt to use the preliminary typology in order to identify outstanding conceptual and
553 practical problems with TIPPME.
554
555 **Methods:** The research team (n=8, excluding the first two listed authors who were responsible for
556 producing the exercise materials), completed a task which encouraged engagement with the detail
557 of the preliminary typology and its application. Each participant was given a link to an online
558 Qualtrics task comprising 40 short (<150 words) intervention descriptions. These represented a
559 sample of descriptions of interventions from the 346 papers that were included in the
560 aforementioned scoping review of choice architecture interventions[23]. We selected intervention
561 descriptions on a quota basis that covered a wide range of intervention content, aiming to include at
562 least 5 examples that could feasibly be mapped to each of the six intervention types within the
563 typology, with a spread across the three intervention foci and across food, alcohol and tobacco. We
564 used the first example that met our criteria that was encountered via random searching to ensure that
565 the intervention descriptions were varied in nature and broadly representative of the wider empirical
566 literature. The 40 intervention descriptions were presented to each participant in a random order.
567 For each intervention example, participants were asked to assess which category in the typology
568 best captured the example, how much overlap existed between the intervention types they
569 considered selecting, and to describe any difficulties they encountered in coding the example and
570 any possible alterations to the typology that would have ameliorated these difficulties. For each
571 intervention example, they were encouraged to provide further qualitative feedback concerning each
572 intervention example, and the overall structure and content of the typology.

573 **Results:** Quantitative and qualitative responses from the task were synthesised. While quantitative
574 results suggested that intervention types could be applied consistently (Fleiss' kappa=.69), the task
575 was principally intended to highlight areas in which there were significant levels of disagreement,
576 comment or criticism, in order to prioritise focused discussion.
577
578 **Subsequent actions:** A teleconference of the research team was convened, and a structured
579 discussion was conducted. This involved assessing problematic intervention descriptions in a
580 structured format, in order to reach agreement on steps to be taken to improve the conceptual
581 coherence and ease of use of the typology. Intervention examples where ≥50% of responses were
582 discordant were flagged for prioritised discussion. Discussion began in order of the intervention
583 descriptions that were coded least consistently, and terminated after all flagged examples had been
584 discussed. For each flagged example, individuals were encouraged to justify their responses and
585 propose and discuss solutions which could overcome the reasons why individuals coded
586 discrepantly. A vote then took place to assess if participants could converge on the same answer i)
587 without any further changes to the typology and ii) with specific changes to the typology (if agreed
588 upon). As a result of this process, various changes were made to wording of definitions, including a
589 clarification of the distinction between the columns within the typology. Additional guidance text
590 was added where it was agreed there was likely to be a greater chance of perceived overlap between
591 intervention types. Figure 3 shows the typology version at this stage.
592
593
594
595
596 **PLEASE INSERT FIGURE 3 HERE**
597
598 **Figure 3. Typology at end of Stage 4**
599
600
601 **Stage 5. Reliability testing exercise (i): Coding of intervention descriptions by external experts**
602

603 **Aim:** To test whether participants likely to use TIPPME, namely those involved in researching or
604 implementing interventions to change health-related behaviours, were consistent in identifying its
605 intervention types and foci in short descriptions of interventions.
606
607 **Methods:** We recruited external experts with backgrounds in public health and behavioural science
608 as researchers or practitioners, meeting the following criteria (adapted from[15]): "active in their field
609 and engaged in designing, delivering and/or evaluating interventions to change health-related
610 behaviour that could be delivered at scale to impact on population health". Recruitment was via
611 email and Twitter enquiries to possible participants within our wider academic networks. A similar
612 exercise to that described in Stage 4 was used, involving coding 40 short intervention descriptions
613 (<150 words) selected on a quota basis to represent a range of intervention content and of targeted
614 products. These were presented in a random order, using the question "Which intervention category
615 best captures the above description?". Quantitative reliability statistics were calculated for the pre-
616 specified primary outcome of discrimination of intervention type (the rows of the typology), as well
617 as for intervention focus (the columns of the typology) and a combined total. As some agreement
618 would be achieved by chance alone, two different 'chance-corrected' agreement measures were
619 used, Fleiss' kappa and prevalence and bias adjusted kappa (PABAK)[42,43].
620
621 **Results:** Potential participants (n=52) who had initially expressed an interest in participating were
622 contacted via email with a link to the exercise. All participants (n=33) who started the exercise
623 completed it. 58% were female, and most were in research roles (94%, with 6% in policy or
624 practitioner roles), with a range of disciplinary backgrounds represented (psychology or behavioural
625 science (45%); public health or medicine (18%); nutrition (15%); epidemiology (6%); economics
626 (6%); other (urban planning, marketing, human factors) (9%)). Inter-rater reliability values for the
627 exercise are provided in Table 2. A kappa value of .41-.60 is conventionally considered to represent
628 "moderate" agreement, a value of 0.61–0.80, "substantial" agreement, and a value of >.80, "almost

perfect" agreement[44]. Fleiss classified a kappa between .40-.75 as "Intermediate to Good" and >.75 as "Excellent"[43]. Therefore, the observed values indicate that the intervention types within the typology in its current form are strongly distinguishable from one another. Furthermore, category-wise statistics indicate that each of the six intervention types and each of the three intervention foci were able to be applied reliably.

Subsequent actions: A two-day residential meeting of the research team was held to further refine the typology. Although the structure of the typology was not altered, each intervention type was given a concise single-word title (e.g. 'Words, numbers and pictures' was changed to 'Information'), and some minor changes to wording of definitions were made. It was agreed that, because intervention descriptions used in this exercise comprised short passages focused on the intervention characteristics, this was not representative of how these might more typically be encountered in full-text papers, where details may be spread disparately within papers, in potentially complex formats. Furthermore, it was important that participants were able to code the presence of multiple discrete intervention types identified within one paper, where previously they had been asked to identify a single category that best captured an intervention. Therefore, a second reliability testing exercise intended to be less artificial and more generalisable to real-world use was planned.

Stage 6. Reliability testing exercise (ii): Coding of intervention descriptions using full-text papers

Aim: To assess whether interventions described in full-text papers can be reliably coded to categories within TIPPME.

654 **Methods:** Following a pilot phase to develop the exercise, members of the research team (n=4) each
655 coded 24 study reports within full-text articles that were randomly selected (using a random number
656 generator) from those included in the initial scoping review[23] but excluding review papers and those
657 not describing any intervention, those concerning physical activity interventions, and those that had
658 been used in previous stages of the typology development process. To ensure that a range of
659 behaviours was covered, randomisation was stratified by behaviour so that half of the papers related
660 to food and half to alcohol or tobacco. Where there were multiple eligible separate studies within a
661 single full-text paper, the first was used. The exercise involved each participant coding which (if
662 any) categories in the typology were identified in each study. The number of study reports to be
663 coded in this exercise was derived using the KappaSize R Package[45], given that, to our knowledge,
664 there are no gold-standard methods to precisely estimate the required sample sizes for determining
665 reliability kappas in cases where there are both multiple coders and a large number of coding
666 categories. We estimated an approximate, conservative sample size based on the following
667 parameters: an alpha value of 0.05; power of 0.80, using 4 coders; an assumption that categories
668 will not be perfectly balanced and instead may be moderately unbalanced; a null hypothesis of a
669 kappa of 0.4 (i.e., the lower bound of 'intermediate to good' agreement on Fleiss' Kappa
670 Benchmark Scale); and, an expected kappa of 0.7 (based on observed kappa values from reliability
671 testing exercise (i)). This suggested that at least 22 study reports would be required to test whether
672 the kappa exceeds 0.4.
673
674 **Results:** Inter-rater reliability values are provided in Table 2. The observed values indicate that the
675 intervention types within the typology in its current form are strongly distinguishable from one
676 another when full-text papers are coded. Furthermore, category-wise statistics indicate that
677 underlying the summary statistics, each of the six intervention types and each of the three
678 intervention foci was able to be applied reliably.
679

680 **Stage 7. Agreement on a final version of TIPPME**

681

682 **Aim:** To reach consensus on a final version of TIPPME including terminology, wording and
683 presentation.

684

685 **Methods:** Research team members completed a final check of the typology and its wording, to
686 ensure it was clear and consistent throughout. A teleconference of the research team was convened
687 to discuss any identified issues.

688

689 **Results:** Further descriptive notes and additional examples were added to the full version of the
690 typology to aid in its use. A final version of TIPPME was agreed upon by the research team,
691 described in the 'Results' section.

692
693

Table 2. Inter-rater reliability statistics (Fleiss' kappa, prevalence and bias adjusted kappa (PABAK), and percent agreement) for reliability testing exercises (i) (left) and (ii) (right)

Reliability testing exercise (i) – external experts (33 coders, 40 codings each)				Reliability testing exercise (ii) – full-text papers (4 coders, 24 codings each)			
Summary statistics				**Summary statistics**			
	Intervention type (primary outcome)	Intervention focus	Total		Intervention type (primary outcome)	Intervention focus	Total
Fleiss' kappa [95% CI]	.76 [.70, .83]	.62 [.53, .72]	.61 [.55, .67]	Fleiss' kappa	.80	.71	.73
PABAK [95% CI]	.77 [.71, .84]	.69 [.62, .76]	.63 [.57, .69]	PABAK	.87	.77	.87
Agreement	.81	.77	.65	Agreement	.93	.88	.94
Category-wise statistics				**Category-wise statistics**			
	Fleiss' kappa [95% CI]	PABAK [95% CI]			Fleiss' kappa [95% CI]	PABAK [95% CI]	
Intervention type				Intervention type			
Availability	.65 [.43, .87]	.85 [.77, .93]		Availability	.68 [.36, 1.00]	.81 [.62, 1.00]	
Position	.93 [.88, .97]	.97 [.95, .99]		Position	1.00 [1.00, 1.00] *	1.00 [1.00, 1.00] *	
Functionality	.67 [.48, .87]	.92 [.86, .98]		Functionality	N/A	N/A	
Presentation	.77 [.68, .87]	.82 [.74, .90]		Presentation	.79 [.61, .98]	.82 [.64, 1.00]	
Size	.75 [.63, .87]	.87 [.79, .95]		Size	.87 [.70, 1.00]	.90 [.76, 1.00]	
Information	.82 [.74, .89]	.86 [.79, .94]		Information	.90 [.77, 1.00]	.92 [.80, 1.00]	
Other	.01 [-.01, .04] *	.94 [.91, .97] *		Other	-.04 [-.09, .00] *	.83 [.67, .99] *	
Intervention focus				Intervention focus			
Product	.65 [.55, .74]	.65 [.55, .74]		Product	.76 [.54, .97]	.76 [.56, .97]	
Related Objects	.51 [.37, .64]	.62 [.52, .72]		Related Objects	.72 [.41, 1.00]	.86 [.70, 1.00]	
Wider Environment	.79 [.67, .91]	.86 [.79, .94]		Wider Environment	.72 [.53, .92]	.74 [.54, .94]	
Other	.01 [-.01, .04] *	.94 [.91, .97] *		Other	-.04 [-.09, .00] *	.83 [.67, .99] *	

* Very few datapoints contained this code (being applied on average less than once per coder over the set of intervention descriptions). Due to its low frequency, this does not allow confidence in associated kappa statistics, which are therefore reported only for completeness.
N/A = Code not used by any coder

SUPPLEMENTARY INFORMATION

Full version of TIPPME and guidance for use.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY

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

REFERENCES

1 GBD 2015 Mortality and Causes of Death Collaborators. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet **388**, 1459-1544 (2016).

2 GBD 2015 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet **388**, 1659-1724 (2016).

3 Thaler, R. H. & Sunstein, C. R. Nudge: Improving decisions about health, wealth, and happiness. (Yale University Press, 2008).

4 Social and Behavioral Sciences Team. Social and Behavioral Sciences Team 2016 Annual Report. Executive Office of the President National Science and Technology Council (2016).

5 Sousa Lourenço, J., Ciriolo, E., Rafael Rodrigues Vieira de Almeida, S. & Troussard, X. Behavioural insights applied to policy: European Report 2016. Publications Office of the European Union (2016).

6 Swinburn, B., Egger, G. & Raza, F. Dissecting Obesogenic Environments: The Development and Application of a Framework for Identifying and Prioritizing Environmental Interventions for Obesity. Preventive Medicine **29**, 563-570 (1999).

7 Kelly, M. P. & Kelly, R. S. in Handbook of Biology and Society (eds M. Meloni, J. Cromby, P. Fitzgerald, & S. Lloyd) (Palgrave Macmillan, 2017).

8 Marteau, T. M., Hollands, G. J. & Fletcher, P. C. Changing Human Behavior to Prevent Disease: The Importance of Targeting Automatic Processes. Science **337**, 1492-1495 (2012).

9 Hollands, G. J., Marteau, T. M. & Fletcher, P. C. Non-conscious processes in changing health-related behaviour: a conceptual analysis and framework. Health Psychology Review **10**, 381-394 (2016).

10 Adams, J., Mytton, O., White, M. & Monsivais, P. Why Are Some Population Interventions for Diet and Obesity More Equitable and Effective Than Others? The Role of Individual Agency. PLOS Medicine **13**, e1001990 (2016).

11 McGill, R. et al. Are interventions to promote healthy eating equally effective for all? Systematic review of socioeconomic inequalities in impact. BMC Public Health **15**, 457 (2015).

12 Hagman, W., Andersson, D., Västfjäll, D. & Tinghög, G. Public Views on Policies Involving Nudges. Review of Philosophy and Psychology **6**, 439-453 (2015).

13 Petrescu, D. C., Hollands, G. J., Couturier, D.-L., Ng, Y.-L. & Marteau, T. M. Public Acceptability in the UK and USA of Nudging to Reduce Obesity: The Example of Reducing Sugar-Sweetened Beverages Consumption. PLOS ONE **11**, e0155995 (2016).

14 Smith, K. E. Beyond Evidence Based Policy in Public Health: The Interplay of Ideas. (Palgrave Macmillan, 2013).

15 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).

16 Kok, G. et al. A taxonomy of behaviour change methods: an Intervention Mapping approach. Health Psychology Review **10**, 297-312 (2016).

17 Johnson, E. J. et al. Beyond nudges: Tools of a choice architecture. Marketing Letters **23**, 487–504 (2012).

18 Münscher, R., Vetter, M. & Scheuerle, T. A Review and Taxonomy of Choice Architecture Techniques. Journal of Behavioral Decision Making **29**, 511-524 (2016).

19 Bitner, M. J. Servicescapes: The impact of physical surroundings on customers and employees. Journal of Marketing **56**, 57-71 (1992).

20 Turley, L. W. & Milliman, R. E. Atmospheric Effects on Shopping Behavior: A Review of the Experimental Evidence. Journal of Business Research **49**, 193-211 (2000).

21 Wadhera, D. & Capaldi-Phillips, E. D. A review of visual cues associated with food acceptance and consumption. Eating Behaviors **15**, 132-143 (2014).

22 Wansink, B. Environmental Factors That Increase the Food Intake and Consumption Volume of Unknowing Consumers. Annual Review of Nutrition **24**, 455-479 (2004).

23 Hollands, G. J. et al. Altering micro-environments to change population health behaviour: towards an evidence base for choice architecture interventions. BMC Public Health **13**, 1218 (2013).

24 Hollands, G. J. et al. Altering choice architecture to change population health behaviour: a large-scale conceptual and empirical scoping review of interventions within micro-environments. University of Cambridge, Cambridge (2013).

564 25 Oliver, A. Nudging, shoving and budging: behavioural economic-informed policy. Public
565 Administration **93**, 700-714 (2015).
566 26 House of Lords Science and Technology Select Committee. Behaviour change: 2nd report of
567 session 2010–12 (HL paper 179). The Stationery Office Limited, London (2011).
568 27 Larsen, K. R. et al. Behavior change interventions: the potential of ontologies for advancing
569 science and practice. Journal of Behavioral Medicine **40**, 6-22 (2017).
570 28 West, R. & Michie, S. A guide to development and evaluation of digital behaviour change
571 interventions in healthcare (version 1). (Silverback Publishing, 2016).
572 29 Cochrane. Cochrane PICO Ontology, <http://linkeddata.cochrane.org/pico-ontology>
573 (2016).
574 30 Dombrowski, S. U., O'Carroll, R. E. & Williams, B. Form of delivery as a key 'active
575 ingredient' in behaviour change interventions. British Journal of Health Psychology **21**,
576 733-740 (2016).
577 31 Arp, R., Smith, B. & Spear, A. D. in Building ontologies with basic formal ontology (eds R.
578 Arp, B. Smith, & A.D. Spear) (The MIT Press, 2015).
579 32 Schutz, A. On Phenomenology and Social relations: Selected Writings. (Chicago University
580 Press, 1970).
581 33 Bucher, T. et al. Nudging consumers towards healthier choices: a systematic review of
582 positional influences on food choice. British Journal of Nutrition **115**, 2252-2263 (2016).
583 34 Hollands, G. J. et al. Portion, package or tableware size for changing selection and
584 consumption of food, alcohol and tobacco. Cochrane Database of Systematic Reviews,
585 CD011045 (2015).
586 35 Hollands, G. J. et al. Altering the availability or proximity of food, alcohol and tobacco
587 products to change their selection and consumption [Protocol]. Cochrane Database of
588 Systematic Reviews **3**, CD012573 (2017).

36 Nudelman, G. & Shiloh, S. Mapping health behaviors: Constructing and validating a common-sense taxonomy of health behaviors. Social Science & Medicine **146**, 1-10 (2015).

37 Hoffmann, T. C. et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ **348**, g1687 (2014).

38 Montgomery, P. et al. Protocol for CONSORT-SPI: an extension for social and psychological interventions. Implementation Science **8**, 99 (2013).

39 Shemilt, I. et al. Pinpointing needles in giant haystacks: use of text mining to reduce impractical screening workload in extremely large scoping reviews. Research Synthesis Methods **5**, 31-49 (2014).

40 Allan, J., Querstret, D., Banas, K. & de Bruin, M. Environmental interventions for altering eating behaviours of employees in the workplace: a systematic review. Obesity Reviews **18**, 214-226 (2017).

41 Arno, A. & Thomas, S. The efficacy of nudge theory strategies in influencing adult dietary behaviour: a systematic review and meta-analysis. BMC Public Health **16**, 676 (2016).

42 Byrt, T., Bishop, J. & Carlin, J. B. Bias, prevalence and kappa. Journal of Clinical Epidemiology **46**, 423-429 (1993).

43 Gwet, K. L. Handbook of Inter-Rater Reliability, 4th Edition: The Definitive Guide to Measuring The Extent of Agreement Among Raters. (Advanced Analytics, LLC, 2014).

44 Landis, J. R. & Koch, G. G. The Measurement of Observer Agreement for Categorical Data. Biometrics **33**, 159-174 (1977).

45 Rotondi, M. A. KappaSize: Sample Size Estimation Functions for Studies of Interobserver Agreement. R package version 1.1 (https://CRAN.R-project.org/package=kappaSize) (2013).

615 **CORRESPONDENCE**

616

617 Correspondence concerning this article should be addressed to Gareth Hollands, Behaviour and
618 Health Research Unit, University of Cambridge, Institute of Public Health, Robinson Way,
619 Cambridge CB2 0SR, UK. E-mail: gareth.hollands@medschl.cam.ac.uk

620

621 **ACKNOWLEDGEMENTS**

622

623 The study was funded by the United Kingdom Department of Health Policy Research Programme
624 (Policy Research Unit in Behaviour and Health [PR-UN-0409-10109]). DO is supported by the
625 Medical Research Council [Unit Programme number MC_UU_12015/6]. The funders had no role in
626 study design, data collection and analysis, decision to publish, or preparation of the manuscript.

627

628 **AUTHOR CONTRIBUTIONS**

629

630 Conceived the study: GJH MPK DO IS SS TMM. Designed and conducted workshops and
631 reliability testing exercises: GJH GB MPK DO IS SS TMM. Conducted and interpreted analysis:
632 All. Preparation of original draft manuscript: GJH GB SS TMM. Drafted and approved final
633 manuscript: All.

634