Health Psychology Review
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rhpr20

Everything should be as simple as possible, but no simpler: towards a protocol for accumulating evidence regarding the active content of health behaviour change interventions

Gjalt-Jorn Ygram Peters[ab], Marijn de Bruin[c] & Rik Crutzen[d]

[a] Department of Methodology & Statistics, Faculty of Psychology, Open University, Heerlen, The Netherlands

[b] Department of Work & Social Psychology, Faculty of Psychology, Maastricht University, Maastricht, The Netherlands

[c] Amsterdam School of Communication Research, Faculty of Society and Behavioural Sciences, University of Amsterdam, Amsterdam, The Netherlands

[d] Department of Health Promotion, CAPHRI, Maastricht University, Maastricht, The Netherlands

Published online: 21 Oct 2013.

To cite this article: Gjalt-Jorn Ygram Peters, Marijn de Bruin & Rik Crutzen (2015) Everything should be as simple as possible, but no simpler: towards a protocol for accumulating evidence regarding the active content of health behaviour change interventions, Health Psychology Review, 9:1, 1-14, DOI: 10.1080/17437199.2013.848409

To link to this article: http://dx.doi.org/10.1080/17437199.2013.848409

Health Psychology Review, 2015
Vol. 9, No. 1, 1–14, http://dx.doi.org/10.1080/17437199.2013.848409

Everything should be as simple as possible, but no simpler: towards a protocol for accumulating evidence regarding the active content of health behaviour change interventions

Gjalt-Jorn Ygram Peters[a,b]*, Marijn de Bruin[c,1] and Rik Crutzen[d]

[a]*Department of Methodology & Statistics, Faculty of Psychology, Open University, Heerlen, The Netherlands;* [b]*Department of Work & Social Psychology, Faculty of Psychology, Maastricht University, Maastricht, The Netherlands;* [c]*Amsterdam School of Communication Research, Faculty of Society and Behavioural Sciences, University of Amsterdam, Amsterdam, The Netherlands;* [d]*Department of Health Promotion, CAPHRI, Maastricht University, Maastricht, The Netherlands*

(*Received 2 May 2013; final version received 22 September 2013*)

There is a need to consolidate the evidence base underlying our toolbox of methods of behaviour change. Recent efforts to this effect have conducted meta-regressions on evaluations of behaviour change interventions, deriving each method's effectiveness from its association to intervention effect size. However, there are a range of issues that raise concern about whether this approach is actually furthering or instead obstructing the advancement of health psychology theories and the quality of health behaviour change interventions. Using examples from theory, the literature and data from previous meta-analyses, these concerns and their implications are explained and illustrated. An iterative protocol for evidence base accumulation is proposed that integrates evidence derived from both experimental and applied behaviour change research, and combines theory development in experimental settings with theory testing in applied real-life settings. As evidence gathered in this manner accumulates, a cumulative science of behaviour change can develop.

Keywords: behaviour change; interventions; methods; techniques; taxonomy; evidence base

Unhealthy diets, binge drinking, exercising insufficiently and smoking are major contributors to disease and premature death, medical expenditures and societal costs related to a loss of labour (Centre for Disease Control, 2012). Health psychology, and more specifically, behaviour change science, can substantially reduce these costs through the development and implementation of (cost-)effective health behaviour change interventions. To this end, over the past decades psychology has accumulated an extensive toolbox of methods of behaviour change. Such methods are psychological principles that can be used to change determinants of behaviour: both 'explicit' psychological variables and processes such as knowledge, risk perception, attitude, anticipated regret, subjective norm and self-efficacy, as well as 'implicit' variables and processes

*Corresponding author. Email: gjalt-jorn@behaviorchange.eu
[1]Current address: Marijn de Bruin, Health Psychology Group, Division of Applied Health Sciences, School of Medicine & Dentistry, University of Aberdeen, Aberdeen, UK

This article was originally published with errors. This version has been corrected. Please see Erratum (http://dx.doi.org/10.1080/17437199.2013.858979).

such as associations and habits (Sheeran, Gollwitzer, & Bargh, 2013). This toolbox of methods of behaviour change has been codified in protocols for intervention development, description and analysis such as Intervention Mapping (IM; Bartholomew, Parcel, & Kok, 1998; Bartholomew, Parcel, Kok, Gottlieb, & Fernández, 2011) and taxonomies for intervention description such as Abraham and Michie's Behaviour Change Technique (BCT) taxonomy (Abraham & Michie, 2008; Michie & Johnston, 2012).

The BCT taxonomies are geared towards facilitating accurate description and coding of intervention content, and they provide a useful overview of BCTs and their definitions. However, the BCT taxonomies do not acknowledge the fact that methods of behaviour change are only effective under specific conditions (Schaalma & Kok, 2009); parameters that are crucial to adequate analysis of intervention content, as well as for designing effective interventions. Although the taxonomy of methods included in the IM protocol (see chapter 6 of Bartholomew et al., 2011) does acknowledge and describe these parameters for the effectiveness of methods of behaviour change, the evidence base for these conditions is suboptimal: no systematic synthesis of the evidence for the parameters for each method is available. Thus, although a versatile and potentially powerful toolbox of behaviour change methods is emerging, there is a strong need for more systematic evidence regarding which methods work and under which conditions.[1]

Recently, many meta-analyses aimed to contribute to such an evidence base (Albarracín et al., 2005; Dombrowski et al., 2012; Johnson et al., 2009; Michie, Abraham, Whittington, McAteer, & Gupta, 2009; Taylor, Conner, & Lawton, 2012; Webb, Joseph, Yardley, & Michie, 2010). These meta-analyses typically examine evaluations of multi-method, applied behaviour change interventions, extract both the reported behaviour change methods and the effect size, and then use meta-regression to examine differences in effect size as a function of whether or not interventions contain a given behaviour change method. While very insightful and useful in their own right, such meta-analyses potentially suffer a number of limitations because in their simplification of reality, they neglect a number of crucial confounders. Specifically, current practice seems to ignore that behaviour change methods have parameters for effectiveness and that methods can interact with each other. Moreover, the influence of contextual factors, such as sample characteristics, behavioural domain and study design, as well as the active content of 'usual' or 'standard care' provided to control groups, is typically not taken into account. We feel that these limitations have not received due attention, which is unfortunate because as a field, we now need to take a step back before we can gain further understanding of behaviour change method effectiveness. In fact, refusing to acknowledge these problems may impede behaviour change science by promoting partial, and thus incorrect, application of psychological theory and disseminating questionable methods for inferring the effectiveness of psychological principles. In the present paper, we will explain our main concerns and suggest possible solutions. Note that our concerns pertain to all meta-analyses employing this methodology of inferring behaviour change method effectiveness from differences in effect size; the papers we cite above are simple illustrations, and by no means particularly flawed.

Methods, applications and conditions for effectiveness

In the discussion that follows, it is crucial to distinguish three closely related concepts. This distinction was proposed in the IM protocol, which also provides useful definitions (Bartholomew et al., 1998, 2011):

(1) Methods of behaviour change: theoretical processes of change according to psychological theory. Behaviour change methods are by definition generic, and not specific to populations or behavioural domains.
(2) Practical applications: the translations of theoretical methods of behaviour change to practical intervention elements. Applications are by definition specific, ideally tailored to populations, intervention contexts and behavioural domains.
(3) Parameters for effectiveness of behaviour change methods: the characteristics that a practical application must manifest for it to accurately reflect the theoretical method. When these parameters are lost in translation from method to application, effective behaviour change is undermined and may even result in counter-productive effects. Evidence for the existence of such parameters can range from theoretical to meta-analytical.

To illustrate these concepts, consider the theoretical *method of behaviour change* of fear appeals, specified in the Extended Parallel Process Model (Witte, 1992). The method is based on the theory that by using fear to target the determinant risk perception (or outcome expectations, included in attitude), behavioural change can be achieved. The *parameters for effectiveness* according to theoretical evidence (Witte, 1992) are that a recipient must (1) perceive a severe threat (severity), (2) consider himself/herself susceptible to this threat (susceptibility), (3) believe that one or several behaviours can effectively diminish the threat (response efficacy) and (4) be confident that he/she can successfully conduct at least one of these behaviours (self-efficacy).[2] Meta-analytical evidence has confirmed that method effectiveness fully depends on perceived threat (conditions 1 and 2 collapsed) and efficacy (conditions 3 and 4 collapsed), and that if the condition of efficacy is not met, the application of fear appeals may even backfire (Peters, Ruiter, & Kok, 2013a). Hence, theory and empirical evidence confirm each other and provide a strong evidence base for when and how to apply fear appeals (and when and how not to apply the method).

Now let us take two intervention developers, each designing an intervention (i.e., comprising one or more *practical applications*) where the target behaviour is engaging in intense physical activity three times a week for at least 30 minutes. The target population of the first developer is 20–30 year-olds who currently engage in physical exercise two times a week. The target population of the second developer is 60–70 year-olds who have been sedentary most of their lives. Most likely, both response- and self-efficacy will be high for the first population, but low for the second population, whereas perceived susceptibility is likely to be low in the first population. This means that the first intervention developer should translate his fear appeal method to an application focusing on susceptibility (note that a focus on severity is rarely wise, see de Hoog, Stroebe, & de Wit, 2007). However, the second developer will need to make sure he or she also increases response- and self-efficacy. Taking into account the target populations, the first intervention developer might create a viral video that uses information from users' social networking sites to enhance susceptibility. The second developer might create an intervention that includes physical exercise sessions to enhance efficacy, combined with quick medical check-ups to enhance susceptibility.

Because a thorough understanding of the concept of parameters for effectiveness is important, we will give another example. The method of behaviour change usually referred to as 'modelling' is based on vicarious learning (Manz & Sims, 1981), and states that behaviour change can be achieved through influencing various determinants (such as

attitude and self-efficacy) through observation of others (models) performing a behaviour. Modelling also has four parameters for effectiveness. The first is shared by many methods of behaviour change: the recipient must attend to the communication and the recipient must remember it, and have sufficient skills to perform the behaviour. In addition, recipients must identify with the model, the model must be positively reinforced for the desirable behaviour and fourth, the model should be a coping model (i.e., realistically struggling with the behaviour) rather than a mastery model (i.e., effortlessly realising change; see Bartholomew et al., 2011). Common mistakes are using a role model with whom people identify poorly (often overlooked when celebrities are used), that performs the right behaviour but is not rewarded in any way (violating the third parameter), or using a model that performs the target behaviour without any problems (violating the fourth parameter).

These descriptions of parameters for effectiveness illustrate three things. First, the effectiveness of a practical application can depend on whether the parameters for effectiveness of the behaviour change method that it embodies are satisfied (e.g., although modelling is an effective method, an application where a celebrity quits smoking instantly and effortlessly is unlikely to contribute to behaviour change). Second, the presence or absence of one accurately applied behaviour change method may contribute to the effectiveness of another (e.g., a successful fear appeal will almost always require a behaviour change method that targets self-efficacy). Third, contextual factors such as behavioural domain and target population can contribute to behaviour change method effectiveness because they can determine whether parameters for effectiveness are satisfied (e.g., some behaviours are easy to perform, but others hard; and some populations are high in self-efficacy, and others low).

The importance of these parameters for effectiveness is not merely a theoretical possibility. The meta-analysis we drew upon earlier provides a clear demonstration of the reality and importance of these conditions (Peters et al., 2013a). It confirmed that fear appeals only work if both perceived threat and perceived efficacy are high. This means that meta-analyses of behaviour change interventions that fail to take these parameters for effectiveness into account risk drawing the wrong conclusions. For example, imagine that such a meta-analysis finds that the effect size of interventions including fear appeals is similar to the effect size of interventions that do not contain this behaviour change method. Its authors might conclude that fear appeals are an ineffective method. However, if the parameters for effectiveness would have been taken into account (which might have revealed that most interventions were conducted in populations low in self-efficacy; or that few of the interventions contained a behaviour change method that successfully enhances self-efficacy), it may turn out that only a small subset of the included interventions applied fear appeals *correctly*, and that in these cases the use of fear appeals did have an added value.

In sum, meta-analyses of behaviour change interventions that do not take into account the parameters for effectiveness run the risk of drawing the wrong or overly simplistic conclusions regarding the effectiveness of behaviour change methods, which in turn may adversely influence intervention developers and practitioners.

Confounding factors in behaviour change meta-analyses

Besides the primary issue of parameters for effectiveness explained above, there are additional concerns regarding evidence on behaviour change method effectiveness

presented by meta-analyses of behaviour change interventions, particularly when they evaluate interventions that test multiple behaviour change methods simultaneously in real-life contexts. This is not to blame the authors – we are amongst them and know that a lot of work and thought goes into such papers – but there is so much potential for interaction between methods of behaviour change and context, that properly controlling for these factors becomes a formidable task.

To illustrate this, we examined data from a meta-analysis of HIV-treatment adherence interventions (de Bruin et al., 2010). In this meta-analysis, detailed information was collected on both the content of intervention care and the content of 'usual' or 'standard' care provided to the control group (for the rationale behind why this is relevant, see de Bruin, Viechtbauer, Hospers, Schaalma, & Kok, 2009; Michie, Prestwich, & de Bruin, 2010). The analyses revealed that there was considerable overlap in the behaviour change methods delivered to intervention and control groups, which strongly increased as the 'control group capacity' increased (the total number of behaviour change methods offered to control groups; $r = .68$, $p < .001$; overlap ranged from 0% to 50% of the behaviour change methods in the intervention manuals). Out of 163 behaviour change methods coded in interventions, 37 (22.7%) overlapped with the care provided to the respective control groups. Some behaviour change methods were 100% unique (every time when they were present in an intervention, they were absent in the control group), whereas others were applied frequently in interventions but never (0%) unique (always also present in the control group). Now, how is it possible to draw conclusions about which method works and which does not, without knowing whether it was really being tested in the first place (i.e., delivered exclusively in the intervention condition)? This issue of control group capacity has been confirmed in subsequent meta-analyses (Freedland, Mohr, Davidson, & Schwartz, 2011; Janssen, de Gucht, Dusseldorp, & Maes, 2013; Waters, St George, Chey, & Bauman, 2012).

Another concern is related to the customary procedure for establishing behaviour change method effectiveness, where the association between method presence and effect size is examined on a method-by-method basis (or sometimes using sets of methods guided by theory, e.g., Michie et al., 2009). This is problematic because some methods may frequently occur simultaneously, warranting measures to control for co-occurrence. For example, when self-monitoring is associated to higher effect sizes, this says nothing about the effectiveness of self-monitoring if self-monitoring is frequently accompanied by, for example, feedback, action planning and tailoring. To check whether method co-occurrence is common, we randomly selected two behaviour change methods using data from the interventions coded in the HIV adherence meta-analysis ($k = 31$; for the coding manual, see http://marijndebruin.eu/meta/hivadherence/taxonomy). Presence of the method 'Provide general information' correlated significantly with presence of 'Social comparison peers' and 'Persuasive communication'; and 'Planning coping responses' even correlated with five other methods (all correlations between 0.36 and 0.51). Thus, in order to examine the effects of a particular behaviour change method, it is necessary to control for method co-occurrence. We are not aware of any meta-analysis that did so based on actual data.

Finally, besides these issues with control group capacity and method co-occurrence, there is the need to control for contextual factors: sample characteristics (e.g., age, socio-economic status), study design choices (e.g., follow-up period, measurement instruments) and study quality indicators (e.g., blinding of study personnel and participants). These can be powerful influences on effects sizes and might correlate with content of interventions (and thus confound the behaviour change method-effect size relationship).

For example, for studies that provided all data including clinical outcomes of patients (k = 15) in the HIV adherence meta-analyses, dropout was strongly related to both the intervention capacity (i.e., the number of included methods) as well as the effect size (r = −0.54 and −0.56). Hence, without controlling these meta-analyses for dropout, we could find an inflated relation between intervention content and effect sizes. More subtle contextual factors can also play a role; for example, some methods might simply be more effective in some cultures than others, for example because determinant relevance may differ (McEachan, Conner, Taylor, & Lawton, 2011). Similarly, it is conceivable that large-scale real-life events can impact method effectiveness; for example, a natural disaster can deplete coping resources, resulting in a low effect size in a given study or perhaps even a set of studies; and it has been found that HIV risk behaviour interventions' effectiveness decreased with epidemic duration (Lacroix, Pellowski, Lennon, & Johnson, 2013).

Acknowledging these confounding roles of control group capacity, method co-occurrence and contextual factors illustrates why we argue that conclusions about the (in) effectiveness of behaviour change methods from meta-analyses of heterogeneous behaviour change interventions are questionable.

The first solution: improving the current practice in behaviour change meta-analyses

There are two ways to take the roles of parameters for effectiveness, method co-occurrence and contextual factors (including control group capacity) into account in meta-analyses of multi-method behaviour change interventions. First, parameters for effectiveness and control group capacity can be taken into account by only including a behaviour change method in the analysis when it has been applied *correctly* (when the parameters for effectiveness hold in the application) and *differentially* (when present in the experimental condition but absent in the control condition). Method co-occurrence and contextual factors would still have to be accounted for in the analysis. Since conclusive evidence on the conditions of effectiveness of most behaviour change methods is as yet unavailable, coding manuals should base their descriptions of conditions for effectiveness on the best available evidence (e.g., theory, combined with data from empirical studies and – if available – reviews; see e.g., the taxonomy in chapter 6 of Bartholomew et al., 2011). A limitation of this approach is that only coding methods when the theoretical parameters for effectiveness are met precludes verification or falsification of these conditions for effectiveness.

A second approach is to take adherence to the presumed parameters for effectiveness, method co-occurrence (i.e., interactions between behaviour change methods) and contextual factors into account by modelling them in the analyses. Note, however, that one method with four parameters for effectiveness would already require four predictors in such a meta-regression model, and modelling parameter adherence for two methods would require eight predictors, even without modelling cross-method interactions between these parameters. A problem of this approach is that usually too few studies are available for the different combinations of these predictors. The supplemental data includes an example of how many dummy variables (103) would have to be included to examine the conditions for only two behaviour change methods (i.e., fear appeal and modelling, which each have four conditions for effectiveness), and this example does not yet take contextual factors into account. Given that meta-analyses generally aim to compare dozens of behaviour change methods, it is clear that this approach is not feasible

(that is, there is a high risk of not having sufficient studies available; Higgins & Thompson, 2004; López-López, Marín-Martínez, Sánchez-Meca, Van den Noortgate, & Viechtbauer, 2013).

An alternative and more pragmatic step forward would be to use dummy variables to distinguish three levels of satisfaction of parameters. For example, when using method absence as reference category, one dummy variable can represent method application consistent with conditions (correct application) and one dummy variable can represent method application that is not or only partially consistent with theoretical conditions (partial application). In this way, also meta-analyses of behaviour change interventions can contribute to the evidence-base on behaviour change methods (i.e., what works and under what parameters?), by verifying whether adherence to sets of parameters as derived from current best evidence indeed results in more effective behaviour change interventions in practice.

Thus, much could already be accomplished by improving the generally applied methodology in meta-analyses of health behaviour change interventions:

(1) Include descriptions of current best evidence regarding parameters in taxonomies used for coding the active content of interventions, enabling the coding of three levels of method application (absence, partial application and consistent application);
(2) Use these descriptions in dummy coding to distinguish these three levels; if this is not feasible, only coding a method when its application is consistent with all of the method's parameters for effectiveness; and conversely, if possible, ideally using dummies to model each individual parameter of each method to test their accuracy;
(3) Acknowledging method co-occurrence (i.e., include co-occurring methods in the model to control for them) and contextual factors (i.e., sample characteristics, behavioural domain, study design choices, risk of bias) by modelling these in the analysis;
(4) Taking into account each study's control group capacity by (1) identifying which behaviour change methods were uniquely delivered to the intervention group (see de Bruin et al., 2009), or (2) including control group content in a bivariate meta-regression model (cf. de Bruin et al., 2010).

Conclusions about the effectiveness of behaviour change methods from published meta-analyses of behaviour change interventions that did not take these issues into account, should be interpreted cautiously; and following these guidelines in future meta-analyses could greatly improve the value of the results obtained through such meta-analyses of behaviour change interventions. Yet, confirming or falsifying individual parameters for effectiveness of each behaviour change method will remain very hard for such meta-analyses, and thus the results of such analyses are unlikely to greatly contribute to theory development. To be able to examine the dynamics of each method's conditions for effectiveness in sufficient depth, we have to turn to another solution.

The second solution: taking one step back

Although this solution is simple, it requires us to take a step back, because in a sense we took two steps forward where we should have taken one. We presently rely on

meta-analyses of multi-method behaviour change interventions to inform us as to the dynamics of our behaviour change theories. However, because these interventions are designed to address real-life problems, they generally utilise a variety of behaviour change methods targeting a host of different determinants. Ideally, from a theoretical viewpoint, we would not have progressed to this stage before developing a sophisticated understanding of the dynamics of the relevant behaviour change methods. Of course, real-life problems such as smoking and insufficient exercise demanded attention, and taking two steps forward by going ahead and employing the then-current behaviour change science knowledge base was justified, incomplete though this knowledge base may have been. However, we now find ourselves in the situation that we need to take a step back to consolidate our position.

In order to acquire a robust evidence base for our toolbox of behaviour change methods, and given the limitations of doing meta-analyses of behaviour change intervention evaluations, it is necessary to incorporate experimental tests in the evidence-building process. Controlled experiments enable manipulation of single behaviour change methods and individual parameters for effectiveness. Various control conditions can be used (lacking manipulation or presenting one or more placebo manipulations) and method co-occurrence can be avoided (except, of course, when this co-occurrence is the matter under investigation). It is important to utilise factorial designs in such experiments, which enable examining both main and interaction effects. Although interactions between parameters for effectiveness, method co-occurrence, or both are important moderators of method effectiveness, full factorial designs remain rare (see e.g., Peters et al., 2013a).[3] Such experiments can be conducted relatively easily and cheaply. For many behaviour change methods, the outcome measures can be the determinants or processes that a method targets rather than behaviour (but not always; for example, a potential for defensive reactions necessitates measuring behaviour change; Peters et al., 2013a). Furthermore, the Internet has become a common medium for behaviour change interventions, and also conducting such experiments online enables relatively easy and cheap data collection.[4] Then, when publishing the results of successful and unsuccessful replications of these experiments, it is important to publish not only the report, but also the study protocols, stimulus materials, questionnaires and computer tasks, data, analysis scripts and output files (Full Disclosure; Crutzen, Peters, & Abraham, 2012; Peters, Abraham, & Crutzen, 2012). Note that both conducting sufficient exact and conceptual replications and Fully Disclosing the relevant resources are essential to enable high-quality meta-analysis of behaviour change method effectiveness and parameters for this effectiveness.

The best of both worlds: an iterative protocol for evidence base accumulation

Ideally, the two types of meta-analyses distinguished in the present paper (i.e., meta-analyses of more fundamental experiments versus meta-analyses of applied health behaviour change interventions) complement each other (see Figure 1). Meta-analyses of applied behaviour change interventions (which optimise external validity) can provide evidence as to which methods of behaviour change may well be effective. When such meta-analyses use the three-level dummy coding scheme we recommended, such meta-analyses can even provide some indications as to whether this evidence is consistent with assumed parameters for effectiveness. Together with the straightforward avenues for research (i.e., garnering evidence as to each method's effectiveness and the veracity of

each method's assumed parameters for effectiveness), the outcomes of such meta-analyses can then set the research agenda for experimental investigations (which optimise internal validity). Meta-analyses of these experiments can then provide more conclusive answers to the question which methods are more and less effective and under which conditions. The role of contextual variables such as sample characteristics and behavioural domain can be explored by including these variables as moderators in such meta-analyses. These meta-analyses thus allow further development of behaviour change theories, which in turn can be used in the development of new behaviour change interventions. Meta-analyses of evaluations of these interventions can then test whether being compliant with the updated parameters for effectiveness does indeed lead to improved intervention effectiveness in real life. This iterative protocol for evidence base accumulation (IPEBA) covers both the development and the testing of theory, and acknowledges the importance of balancing internal and external validity in the process of building a toolbox of behaviour change methods. While the IPEBA is summarised in Figure 1, Table 1 summarises the underlying limitations of each of the four study types that together necessitate an iterative approach combining all four types, as well as how the combination of these studies in the IPEBA can address these limitations.

IPEBA in real life

When comparing the IPEBA as we presently outlined it to reality, two challenges become apparent. We will briefly discuss those challenges.

First, although from a purely scientific point of view, factorial designs where a small number of parameters for effectiveness and/or methods for behaviour change are systematically varied are preferable, in reality, funders often demand effective interventions that address real-life problems. In such situations, researchers are under considerable

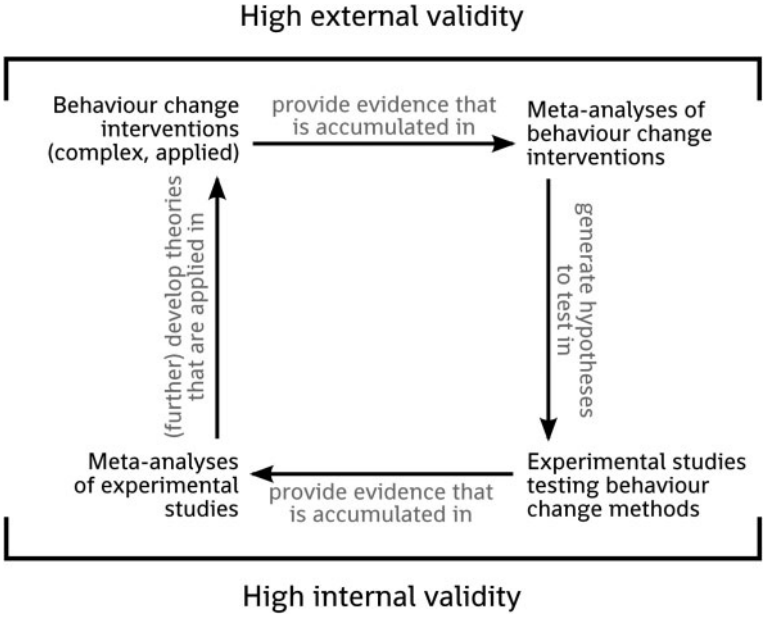

Figure 1. An iterative protocol for evidence base accumulation.

Table 1. Summary of the proposed iterative protocol for evidence base accumulation. The first four rows (1.1–1.4) represent four consecutive sets of studies. After such a sequence, the cycle repeats with 2.1, a reiteration of 1.1, etc.

	Study type	Conditions for effectiveness…	Co-occurrence of methods…	Contextual factors (e.g., sample characteristics, behavioural domain, study design)…
↓	1.1: Effectiveness evaluation of applied behaviour change intervention	…influence the impact of methods on intervention effectiveness and are variably applied …	…can introduce moderation effects and is in the nature of applied interventions…	… influence intervention effectiveness and are a given in real life situations…
RESEARCH CYCLE	1.2: Meta-analysis of effectiveness evaluations of applied behaviour change interventions	… and are hard to take into account in meta-analyses due to complexity…	… and is hard to take into account in meta-analyses due to complexity…	… and can and should be identified and modelled in meta-regressions of intervention evaluations; …
	1.3: Experimental study of behaviour change method and conditions	… but can be manipulated in a full factorial design…	… but can be manipulated in a full factorial design…	… influence method effectiveness even in experimental studies (e.g., when unknown or when they cannot be manipulated) …
	1.4: Meta-analysis of experimental studies of behaviour change method and conditions	… and this evidence can be integrated in meta-regressions of experimental studies to advance behaviour change theory.	… and this evidence can be integrated in meta-regressions of experimental studies to advance behaviour change theory.	… and can and should also be identified and modelled in meta-regressions of experimental studies.
	2.1: Effectiveness evaluation of applied behaviour change intervention	The improved behaviour change theory can then be applied in applied behaviour change interventions…		Applied behaviour change interventions can be better tailored to contextual factors…
↓	2.2: Meta-analysis of effectiveness evaluations of applied behaviour change interventions	…that can be meta-analysed to assess external validity of behaviour change theory and formulate additional hypotheses.		… and whether this tailoring indeed improves effectiveness can then be confirmed through meta-analysis.
	2.3: Experimental study of … etc.	…		…

pressure to deliver an intervention that influences whatever are the outcomes of interest. However, such situations need not be inconsistent with the IPEBA; interventions that are optimised for effectiveness can still contribute useful information, if they are described in sufficient detail. After all, meta-analyses of such behaviour change interventions (see top half of Figure 1) are the only way to test the external validity of the conclusions from meta-analyses of more controlled experiments (see 2.2 in Table 1). This does, however, require very detailed description of the intervention contents. Preferably, all relevant materials and protocols are Fully Disclosed with the study (Peters et al., 2012). The fact that many evaluations that have been published as yet did not Fully Disclose unfortunately means that presently, meta-analyses that try to adhere to the standards implied by the IPEBA will likely only be able to include few studies. Improving the quality of our reporting is as important as prudently approaching meta-analysis of our evidence. Note that Full Disclosure of an intervention is not limited to providing the intervention protocols that describe the intended intervention delivery; preferably, an accurate description of the actual intervention delivery is provided (e.g., as registered through process evaluation or video materials).

Second, given this somewhat unfortunate current state of the literature, what to do when in a meta-analysis, the combinations of parameters of effectiveness that are necessary to draw conclusions are not present in the literature? In this case, we suggest that the meta-analysts make this lack of evidence very explicit. In our view, this serves two goals. First, intervention developers need to know that evidence is lacking, because this (1) renders pilot studies of interventions containing such methods even more vital than they already are and (2) necessitates thorough process evaluation of the effectiveness of the relevant method in influencing the determinant it targets (i.e., intervention evaluations should measure to which degree a method influences the targeted determinants). Second, this lack of evidence necessitates allocation of resources to conduct studies to provide this evidence; thus, funders need to become aware of these lacunae. There is no way around the crucial role of replications (both exact and conceptual, i.e., in different contexts, with different populations, different behaviours, etc.) in advancing psychological theory. This should become routine practice for all our methods of behaviour change.

In addition to these challenges, we would like to emphasise that this paper aims to outline a general protocol that enables establishing the effectiveness of behaviour change methods and studying the parameters of effectiveness that govern each method's effectiveness. This paper does not aim to cover all aspects that are important to consider when doing meta-analyses of behaviour change interventions, such as for example the quality of intervention delivery (when meta-analysists code method presence on the basis of methods sections or intervention protocols, such methods may in practice have been delivered differently or not at all), dose-response effects (a method's effectiveness may be a function of its dose in a linear or nonlinear fashion, see e.g., de Bruin et al., 2009, for an example of dose coding) and the nature of the comparison between the intervention and control conditions (e.g., did the control group receive usual care, a placebo or were they a waiting list control group?). Each of these issues is important and deserves attention, but all fall without the scope of the present paper.

Conclusion

Note that we are aware that many of the limitations we raised are not new (although they have not been made explicit in the literature in this context); and in fact, we have also

discussed several of these with authors of the taxonomies and of the meta-analyses cited here. However, we felt it was time to combine these points and make clear that meta-analyses of health behaviour change interventions on their own cannot provide the kind of evidence that behaviour change science needs; and that intervention developers should pay attention as to whether taxonomies include current best evidence on conditions of behaviour change method effectiveness before they use it as a tool for intervention development. On the other hand, the field of health psychology has the obligation to provide clear guidelines to intervention developers. It has been shown that intervention developers, policy-makers and politicians are insufficiently aware of the complications of behaviour change (Michie & West, 2013; Peters, Ruiter, & Kok, 2013b), so it seems prudent to use IPEBA to develop clear guidelines for intervention development.

By striving to formulate sets of hypotheses in each step of the process, verification and falsification of hypothesised conditions for effectiveness can occur swiftly and accelerate evidence-accumulation (cf. de Bruin & Johnston, 2012; Platt, 1964). Figure 1 and Table 1 show that as evidence accumulates from meta-analyses of controlled experiments, the thus improved behaviour change theory can then be applied in applied interventions that can be meta-analysed to assess external validity. By applying this iterative protocol, we can truly develop a cumulative science of behaviour change and confidently step forward.

Supplemental data

Supplemental data for this article can be accessed here: 10.1080/17437199.2013.848409

Notes

1. Note that we use methods to refer to the methods of behaviour change as defined within IM, and to distinguish these from BCTs. We chose to adhere to IM terminology because these definitions acknowledge (and thus provide definitions and a vocabulary for) the theoretical parameters for effectiveness and the distinction between methods and applications. Both are crucial to the concerns set forth in the present paper.
2. Note that an additional condition for effectiveness is of course whether the targeted determinant predicts the relevant behaviour. For example, even when a method to enhance self-efficacy for chlamydia testing is applied perfectly, it will still not be effective in a population with low perceived susceptibility, as self-efficacy does not predict chlamydia testing among persons who are convinced they are not at risk of having chlamydia. This fundamental condition of effectiveness holds for all methods of behaviour change: for a method to be effective, it must target a determinant it is able to change (e.g., guided practice cannot change subjective norms), and the targeted determinant must predict the relevant behaviour.
3. Note that full factorial designs are desirable for evaluations of interventions as well, but they are much harder to implement in such applied settings.
4. For example using free open source software such as LimeSurvey (LimeSurvey Project Team/ Carsten Schmitz, 2012) or OpenSesame (Mathôt, Schreij, & Theeuwes, 2012).

References

Abraham, C., & Michie, S. (2008). A taxonomy of behavior change techniques used in interventions. *Health Psychology*, *27*, 379–387. doi:10.1037/0278-6133.27.3.379

Albarracín, D., Gillette, J. C., Earl, A. N., Glasman, L. R., Durantini, M. R., & Ho, M.-H. (2005). A test of major assumptions about behavior change: A comprehensive look at the effects of passive and active HIV-prevention interventions since the beginning of the epidemic. *Psychological Bulletin*, *131*, 856–897. doi:10.1037/0033-2909.131.6.856

Bartholomew, L. K., Parcel, G. S., & Kok, G. (1998). Intervention mapping: A process for developing theory- and evidence-based health education programs. *Health Education and Behavior*, *25*, 545–563. doi:10.1177/109019819802500502

Bartholomew, L. K., Parcel, G. S., Kok, G., Gottlieb, N. H., & Fernández, M. E. (2011). *Planning health promotion programs: An intervention mapping approach* (3rd ed.). San Francisco, CA: Jossey-Bass.

Centre for Disease Control. (2012). *CDC – Chronic Disease – Overview*. Retrieved from http://www.cdc.gov/chronicdisease/overview/index.htm

Crutzen, R., Peters, G.-J. Y., & Abraham, C. (2012). What about trialists sharing other study materials? *BMJ*, *345*, e8352–e8352. doi:10.1136/bmj.e8352

De Bruin, M., & Johnston, M. (2012). Methods in health psychology: How do we know what we really know? *The European Health Psychologist*, *14*(4), 107–112.

De Bruin, M., Viechtbauer, W., Hospers, H. J., Schaalma, H. P., & Kok, G. (2009). Standard care quality determines treatment outcomes in control groups of HAART-adherence intervention studies: Implications for the interpretation and comparison of intervention effects. *Health Psychology*, *28*(6), 668–674. doi:10.1037/a0015989

De Bruin, M., Viechtbauer, W., Schaalma, H. P., Kok, G., Abraham, C., & Hospers, H. J. (2010). Standard care impact on effects of highly active antiretroviral therapy adherence interventions: A meta-analysis of randomized controlled trials. *Archives of Internal Medicine*, *170*, 240–250. doi:10.1001/archinternmed.2009.536

De Hoog, N., Stroebe, W., & de Wit, J. B. F. (2007). The impact of vulnerability to and severity of a health risk on processing and acceptance of fear-arousing communications: A meta-analysis. *Review of General Psychology*, *11*, 258–285. doi:10.1037/1089-2680.11.3.258

Dombrowski, S. U., Sniehotta, F. F., Avenell, A., Johnston, M., MacLennan, G., & Ara ujo-Soares, V. (2012). Identifying active ingredients in complex behavioural interventions for obese adults with obesity-related co- morbidities or additional risk factors for co-morbidities: A systematic review. *Health Psychology Review*, *6*, 7–32. doi:10.1080/17437199.2010.513298

Freedland, K. E., Mohr, D. C., Davidson, K. W., & Schwartz, J. E. (2011). Usual and unusual care: Existing practice control groups in randomized controlled trials of behavioral interventions. *Psychosomatic Medicine*, *73*, 323–335. doi:10.1097/PSY.0b013e318218e1fb

Higgins, J. P. T., & Thompson, S. G. (2004). Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine*, *23*, 1663–1682. doi:10.1002/sim.1752

Janssen, V., de Gucht, V., Dusseldorp, E., & Maes, S. (2013). Lifestyle modification programmes for patients with coronary heart disease: A systematic review and meta-analysis of randomized controlled trials. *European Journal of Preventive Cardiology*, *20*, 620–640. doi:10.1177/2047487312462824

Johnson, B. T., Scott-Sheldon, L. A. J., Smoak, N. D., Lacroix, J. M., Anderson, J. R., & Carey, M. P. (2009). Behavioral interventions for African-Americans to reduce sexual risk of HIV: A meta-analysis of randomized controlled trials. *Journal of Acquired Immune Deficiency Syndrome*, *51*, 492–501. doi:10.1097/QAI.0b013e3181a28121

Lacroix, J. M., Pellowski, J. A., Lennon, C. A., & Johnson, B. T. (2013). Behavioural interventions to reduce sexual risk for HIV in heterosexual couples: A meta-analysis. *Sexually Transmitted Infections*. doi:10.1136/sextrans-2013-051135

LimeSurvey Project Team / Carsten Schmitz. (2012). *LimeSurvey: An open source survey tool*. Hamburg. Retrieved from http://limesurvey.org

López-López, J. A., Marín-Martínez, F., Sánchez-Meca, J., Van den Noortgate, W., & Viechtbauer, W. (2013). Estimation of the predictive power of the model in mixed-effects meta-regression: A simulation study. *The British Journal of Mathematical and Statistical Psychology*. doi:10.1111/bmsp.12002

Manz, C. C., & Sims, H. P. (1981). Vicarious learning: The influence of modeling on organizational behavior. *The Academy of Management Review*, *6*(1), 105–113.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*, 314–324. doi:10.3758/s13428-011-0168-7

McEachan, R. R. C., Conner, M., Taylor, N. J., & Lawton, R. J. (2011). Prospective prediction of health-related behaviours with the theory of planned behaviour: A meta-analysis. *Health Psychology Review*, *5*, 97–144. doi:10.1080/17437199.2010.521684

Michie, S., Abraham, C., Whittington, C., McAteer, J., & Gupta, S. (2009). Effective techniques in healthy eating and physical activity interventions: A meta-regression. *Health Psychology*, *28*, 690–701. doi:10.1037/a0016136

Michie, S., & Johnston, M. (2012). Theories and techniques of behaviour change: Developing a cumulative science of behaviour change. *Health Psychology Review*, *6*, 1–6. doi:10.1080/17437199.2012.654964

Michie, S., Prestwich, A., & de Bruin, M. (2010). Importance of the nature of comparison conditions for testing theory-based interventions: Reply. *Health Psychology*, *29*, 468–470. doi:10.1037/a0020844

Michie, S., & West, R. (2013). Behaviour change theory and evidence: A presentation to government. *Health Psychology Review*, *7*, 1–22. doi:10.1080/17437199.2011.64449445

Peters, G.-J. Y., Abraham, C. S., & Crutzen, R. (2012). Full disclosure: Doing behavioural science necessitates sharing. *The European Health Psychologist*, *14*(4), 77–84.

Peters, G.-J. Y., Ruiter, R. A. C., & Kok, G. (2013a). Threatening communication: A critical re-analysis and a revised meta-analytic test of fear appeal theory. *Health Psychology Review*, *7* (Suppl. 1), S8–S31. doi:10.1080/17437199.2012.703527

Peters, G.-J. Y., Ruiter, R. A. C., & Kok, G. (2013b). Threatening communication: A qualitative study of fear appeal effectiveness beliefs among interventionists, policymakers, politicians, scientists and advertisers. *International Journal of Psychology.* doi:10.1002/ijop.12000

Platt, J. R. (1964). Strong inference. *Science*, *146*, 347–353. doi:10.1126/science.146.3642.347

Schaalma, H., & Kok, G. (2009). Decoding health education interventions: The times are a-changing'. *Psychology & Health*, *24*, 5–9. doi:10.1080/08870440801995802

Sheeran, P., Gollwitzer, P. M., & Bargh, J. A. (2013). Nonconscious processes and health. *Health Psychology*, *35*(5), 460–473. doi:10.1037/a0029203

Taylor, N., Conner, M., & Lawton, R. (2012). The impact of theory on the effectiveness of worksite physical activity interventions: A meta-analysis and meta-regression. *Health Psychology Review*, *6*, 33–73. doi:10.1080/17437199.2010.533441

Waters, L., St George, A., Chey, T., & Bauman, A. (2012). Weight change in control group participants in behavioural weight loss interventions: A systematic review and meta-regression study. *BMC Medical Research Methodology*, *12*(1), 120. doi:10.1186/1471-2458-10-110

Webb, T. L., Joseph, J., Yardley, L., & Michie, S. (2010). Using the internet to promote health behavior change: A systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *Journal of Medical Internet Research*, *12*(1), e4. doi:10.2196/jmir.1376

Witte, K. (1992). Putting the fear back into fear appeals: The extended parallel process model. *Communication Monographs*, *59*, 329–349. doi:10.1080/03637759209376276