Complex interventions

Geoff Bates

Institute of Policy Research, University of Bath, Bath, UK

ABSTRACT

Methods of analysis in systematic reviews of interventions continue to develop as authors seek to answer questions about not only whether interventions work, but how and why they work. As our knowledge improves about how complexity in intervention design, implementation and context explains the variation in the association between interventions and outcomes, systematic reviewers increasingly seek to understand this complexity. This article explores how authors of systematic reviews can through their analysis provide the evidence that supports our understanding of how sport and exercise psychology interventions are intended to work and how complexity in the interventions themselves and in the environment they are introduced into may explain the variation in their impacts. It identifies key methods and tools to support reviewers to understand and analyse complexity in interventions, illustrated with examples from sport and exercise psychology. While complexity can be challenging to incorporate into an analysis of interventions in a systematic review, it is an important step for reviewers to take to make sense of large and complex bodies of evidence and to inform the efficient development and adaptation of interventions.

ARTICLE HISTORY

Received 1 March 2021
Accepted 25 October 2021

KEYWORDS

Systematic review; evidence synthesis; complex interventions; complexity; behaviour change

Introduction

This article explores how authors of systematic reviews (hereinafter referred to as ‘reviewers’) can improve the analysis of complexity in systematic reviews of interventions. Sport and exercise psychology interventions delivered in a natural or ‘real world’ context that aim to prevent, promote, or maintain outcomes or behaviours of interest are often described as being ‘complex interventions’. The definition of what makes an intervention complex is debated in the literature, but typically includes that interventions have multiple interacting components and try to address multifaceted issues, and contrast with medical interventions, such as a treatment procedure or medication, through their non-linearity (Petticrew, 2011). Complexity goes beyond the multifaceted nature of the intervention itself into how they are implemented in the messy ‘real world’ context that they are delivered in. Many factors influence how an intervention is delivered and the impacts it has and it is suggested that the association between exposure to an
intervention and an outcome is influenced by three key areas of complexity: (i) intervention design, (ii) environmental context, and (iii) the implementation of the intervention (Anderson et al., 2013). For example, in a recent example of a mindfulness-based mental health intervention targeting student athletes (Hut et al., 2021) the authors suggest that fewer than half participants from track and field teams attended all six sessions that formed the intervention, in comparison to athletes in other sports who attended all sessions. This reduced attendance might suggest that there is something that limited participation from track and field athletes specifically. This could be something in the design or delivery of the intervention, or something about these athletes or their environment, such as their schedules, group norms, or attitudes that influenced their ability to attend and therefore, potentially, limited the effectiveness of the intervention.

Rather than thinking of interventions as simply complex or not, they can be conceptualised as falling along a spectrum of complexity. An Intervention Complexity Assessment Tool (iCAT_SR) has been developed to help reviewers assess where interventions sit on this spectrum and to support their analysis (Lewin et al., 2017). Complexity presents challenges to reviewers such as in defining the interventions to be studied, developing standardised inclusion criteria, and synthesising data from a range of interventions with high heterogeneity (Shepperd et al., 2009). Even when reviewers include interventions that are broadly similar in goal, setting, or populations they are still likely to have many differences (Higgins et al., 2019). The mindfulness intervention in the example above shares clear similarities with a second recently evaluated mindfulness-based intervention that sought to improve mental health in student athletes (Shannon et al., 2018). However, there are differences apparent between these interventions in their details. For example, one intervention was heavily influenced by Self Determination Theory and included a workshop followed by a two week period of self-directed app-based sessions (Shannon et al., 2018), compared to the delivery of six weekly group sessions (Hut et al., 2021). They were also set in different countries and university systems.

In addition to focusing on the overall effects of interventions, reviewers are therefore also interested in exploring whether these distinguishing factors mediate these effects. Rather than asking whether mindfulness interventions support good mental health in student athletes, we might want to know how the use of theory in these interventions mediates their impacts, how app-based sessions compare to multiple group sessions, or how differing norms and values relating to mental health might mediate outcomes. Systematic reviews are a valued and efficient method for answering questions about interventions, traditionally whether they are effective but increasingly how and why this variation and complexity in their design, implementation and context might explain any variation in effectiveness. Systematic reviews aim to synthesise data from primary research studies that meet pre-determined criteria to answer specific research questions (Chandler et al., 2021). Analysis is therefore at the heart of any systematic review and represents the point where reviewers examine the data they have extracted from different studies in an attempt to answer their research questions. Through their analysis reviewers identify patterns in the data through which they will draw conclusions, and make recommendations for policy, practice, and research.

There are many methods for analysing data in a systematic review and the approach taken needs to be appropriate for the data extracted, which in turn must reflect the
aims of the review (Tricco et al., 2011). For example, a thematic synthesis of qualitative studies is appropriate to understand older people’s perceptions about physical activity (Franco et al., 2015), but to identify how effective different interventions have been at enhancing activity amongst older people reviewers might choose to look at a synthesis of quantitative data (Kwan et al., 2020). These reviews utilise very different methods of analysis to answer just two of the many types of questions that systematic reviews seek to answer. It is not the intention here to try and describe all the interesting analysis methods that reviewers can draw on. There are excellent guidance and methodological papers that provide rigorous overviews for methods to synthesise different types of data (Campbell et al., 2020; Higgins et al., 2019; Lau et al., 1997; Thomas & Harden, 2008; Wong et al., 2014). A first takeaway point however is that reviewers should not start off by selecting the method of analysis, but instead, it should be chosen to enable them to answer their review questions and provide findings that are useful to their intended audience.

Where reviewers ask complex questions about interventions then they need to draw upon appropriate methods to answer these questions, as traditional methods often encourage the simplification of this complexity and use linearity to frame interventions and their impacts (Anderson et al., 2013). This article explores the importance of analysing complexity of interventions in systematic reviews of sport and exercise psychology interventions and identifies methods and tools designed to support reviewers with this analysis. Part one of this article starts with a broad discussion of the analysis of interventions in a systematic review and an overview of common methods. It then focuses on approaches to examine complexity of interventions in their design, implementation, and environmental context. The emphasis is on the importance of understanding not just whether interventions are effective or not, but why and how this is the case and on improving the usefulness of the evidence that systematic reviews provide. Part two provides guidance to support analysis of interventions in systematic reviews and identifies methods and tools to support reviewers to understand and analyse complexity, and to report their findings.

**Part 1: Analysing interventions in systematic reviews**

The term intervention is used to cover a substantial variety of approaches that are designed to bring about some change in response to an identified need or problem. Researchers and practitioners use a range of terms with little consistency to categorise interventions such as programme, treatment, therapy, event, initiative, or training. Interventions can have many different purposes and implications, such as to treat an existing problem or to prevent a future problem from happening. They may aim to achieve broad and long-lasting changes or very specific and time-limited goals, and target individuals, groups, organisations, communities or societies. Therefore, they vary hugely in their ambition, design, implementation and complexity, as different methods are naturally required to bring about different types of changes (Kok et al., 2016).

Many different interventions are developed and tested in different contexts in response to a broadly similar problem, perhaps due to different interpretations of that problem or expectations of what is required to solve it. The effectiveness of these interventions varies and this can contribute to a potentially large and complex evidence base,
which can be challenging to make sense of or to identify examples of best practice. Following the principles of evidence-based medicine, with its emphasis on identifying and applying the best relevant evidence (Sackett et al., 1996), systematic reviews are understood to help us identify which interventions are likely to be most effective at responding to a problem. Substantial advances have been made towards building a science of interventions in the past two decades in developing methods, frameworks, models, and approaches to design, evaluate and implement interventions (Craig et al., 2008; Michie et al., 2014; Minary et al., 2019; Movsisyan et al., 2019; Nilsen, 2015; O’Cathain et al., 2019; Richards & Hallberg, 2015). These are designed primarily to support the development of interventions that are robustly designed and built upon a thorough understanding of the problem they seek to address. These methods have emerged mainly from the health and behavioural sciences, and ultimately aim to support the development of interventions that are cost-effective, sustainable and replicable, and do not cause harm. As those developing and evaluating interventions have explored these issues, so too have those who wish to synthesise the findings from studies that report on them. Many of the ideas and tools that have been designed to improve this science of interventions can also be used by reviewers to analyse interventions.

**Common approaches to synthesising evidence from studies of interventions**

When considering how to analyse data in a systematic review of interventions, reviewers likely think firstly of meta-analysis. Indeed, it would be impossible to present a discussion on analysis without highlighting its significance. Where possible to do so in systematic reviews of interventions, data is typically synthesised in a meta-analysis in order to calculate an overall effect size that indicates the strength and direction of the association between an intervention and outcome (Munn et al., 2014). It provides a statistical method that enables reviewers to combine results from what can be very large and messy datasets, which may appear to contain conflicting evidence on intervention effectiveness, into one estimation of the effect of an intervention. This improves precision of the estimation of the intervention effect, to answer questions not asked by the individual studies included in a review, and to assess and understand the importance of differences within conflicting results across studies (Deeks et al., 2019).

Systematic review and meta-analysis are terms often used interchangeably, which can (falsely) give the impression that they are the same method. A meta-analysis has the specific and clear purpose to determine the size and direction of the effect of intervention by combining statistical results from two or more studies (Deeks et al., 2019). Synthesising data in this way helps reviewers to be more confident about the association between interventions and outcomes, and provides a more accurate estimate of this relationship than would be possible from presenting the findings of the studies separately. However, we should not make the mistake of considering meta-analyses and systematic reviews to be one of the same. Firstly, statistical combining of data from different studies is just one method of analysis and one that is not always appropriate. Typically, this is when the data is insufficient or represents a heterogenous selection of studies that if combined is likely to produce misleading or unhelpful results (Deeks et al., 2019). Secondly, a meta-analysis can be used to analyse data across studies where the researchers have not necessarily followed the processes needed to identify and select studies required in a
systematic review (Munn et al., 2014). Meta-analyses based upon non-systematic review methods of identifying and assessing studies however are at risk of bias, and findings from such studies are less reliable (Tod & Eubank, 2017). Finally, where reviewers are interested in exploring complexity in a systematic review their analysis needs to go beyond calculating intervention effect sizes and therefore while it may include a meta-analysis where appropriate to do so, additional methods of analysis are needed.

A second common analysis approach in a systematic review of quantitative evaluations of interventions is to undertake a narrative synthesis. This should not be confused with a ‘narrative review’, a term that is commonly used to refer to ‘non-systematic’ literature reviews. It also differs from a type of systematic review known as a meta-narrative review, which follows a method designed to examine how a topic has been conceptualised and researched by groups of researchers at different times or within different disciplines (Wong et al., 2013). The simple explanation of a narrative synthesis in a review of interventions is that reviewers bring together and make sense of the findings from different studies by describing the intervention effects in passages of text. Popay and colleagues (2006) liken this process to telling stories in a transparent and rigorous way, supported by presentations of data in tables and figures so that it is easily accessible and understandable to the reader. While narrative synthesis is sometimes suggested to be what reviewers resort to when they cannot undertake a meta-analysis, leaving the impression that it is a weaker or less desirable analysis approach, it has its own strengths. One of the main purposes of producing a systematic review is to make sense of the evidence so as to make clear recommendations for policy and practice to decision-makers. Explaining the findings using passages of text rather than presenting solely statistical summaries supports reviewers to make a convincing case (Popay et al., 2006). Additionally, in contrast to a meta-analysis, reviewers can more easily incorporate different study designs and diverse sources of evidence in a narrative synthesis.

Narrative synthesis can be understood therefore as a method with a purpose of its own. In reviews of interventions, it can complement a meta-analysis where it is possible to undertake one that incorporates some or all of the included studies, or provide a useful standalone analysis where a meta-analysis is either not appropriate or needed. One reason why narrative synthesis is perhaps seen as the weaker relation of meta-analysis is that it has been criticised as introducing bias through a lack of transparency in how findings relate to the data (Campbell et al., 2019). To address this problem, newly prepared guidance aims to support reviewers to undertake narrative synthesis in quantitative reviews of interventions without losing the rigour, transparency and quality associated with meta-analysis. The SWiM (Synthesis Without Meta-analysis) reporting standards (Campbell et al., 2020) are a helpful resource for reviewers that encourages transparency in reporting by emphasising the need to describe and justify decisions made in applying the methodology.

Regardless of whether reviewers synthesise data on the effectiveness of interventions using statistical or narrative methods, or both, analysis that only examines this question of effectiveness are narrow in scope and only tell decision-makers some of what they need to know (Snilstveit et al., 2012). This is not to dismiss the clear importance of identifying intervention effectiveness. Put simply, reviewers will always be interested in identifying the extent to which one intervention (or type of intervention) works in comparison to another and whether it brings about the changes in behaviour or outcomes that have
been identified as important. However, readers are often not only interested in whether an intervention is effective or not, but why this is the case. Therefore, the straightforward question of what effect interventions have can be considered the start of a process in a systematic review that progresses to explore the similarities and differences between them (Petticrew et al., 2013). Additional analysis beyond that which helps us to understand if interventions work or not is required to help reviewers understand the direction of the effect, including through analysing complexity in the design, implementation and environmental context of interventions. Here we discuss methods of analysis that reviewers can use to explore complexity in these three areas.

**Complexity in intervention design**

One key area of complexity that reviewers have sought to engage with is the complexity of the intervention itself. Physical activity promotion is a good example of a field of interventions with common aims that have been based on different theoretical concepts and a range of intervention designs. This creates a complicated and messy evidence base for anyone hoping to identify what interventions should be developed or adapted to enhance physical activity rates in their own context. For example, the effectiveness of school-based interventions to increase physical activity amongst adolescents appears to vary greatly when findings are synthesised and there is debate about which intervention designs should be pursued (Kriemler et al., 2011). Evaluations of these school-based physical activity interventions reveal substantial variation in impact, and this may be explained in a systematic review by a wide range of factors, such as: how the intervention is delivered; who provides it; and its content or length (Hynynen et al., 2016). A task for reviewers is therefore to not only identify whether a broad approach (for example, a classroom-based programme) appears to be effective, but to explain the variation in effectiveness across data from studies of these broadly similar approaches. In the context of these school-based physical activity interventions, to make recommendations about how to increase physical activity in adolescents reviewers, therefore, need to analyse the design of these interventions and understand what it is about them that means they bring about change or not, and to what extent. Two ways to explore how an intervention is intended to work in our analysis is to identify its content and its theoretical basis.

**Analysing intervention content**

Interventions vary greatly in content and design and in how they attempt to bring about change. Complex interventions therefore often include multiple components. Reviews can scrutinise the components of interventions by breaking them down and examining the specifics of what they include. This helps to provide greater insights into the nature of the evidence base and to help understand the differences in the effects that these interventions have on their outcomes of interest. For example, scrutiny of a recent and well-reported intervention designed to integrate high-intensity interval training in the workplace (Eather et al., 2020) reveals that the intervention included various components designed to maximise participation including choice of workout session type, reminders for non-attendance, individual or partner-based workouts, real-time feedback of physiological measures, and monitoring of adherence to the programme. Identifying the specific components of complex interventions such as this gives us greater
insight into how that intervention was intended to work and, when reviewers synthesise evidence from multiple interventions, this helps them to understand whether design features are collectively associated with increased effectiveness. In this example, analysis could indicate that interventions that offer choice of workout session as a common component appear more effective, which those seeking to developing workplace interventions can learn from.

Being as specific as possible about the content of interventions included in a review therefore helps the reader to understand the evidence being presented, and to adapt an intervention where desired. However, while there are clearly many benefits of scrutinising intervention design in this manner, there is a risk that by attempting to break down interventions into the smallest components that reviewers can lose sight of the bigger picture and of the intervention as a whole. There is balance needed between wanting to make sense of interventions by analysing these specific components, while maintaining a focus on the context that they are delivered in, and reviewers should be mindful that it may be through the combined effect of components within a complex intervention that is designed to be effective rather than individual components. Building on the previous example, it may be that offering choice of workout session is most effective when an intervention also includes real-time feedback on physiological measures.

In our review of anabolic androgenic steroid (AAS) prevention interventions (Bates et al., 2019), we benefited from identifying the specific components in intervention content and assessing how interventions were expected to work. On a superficial level, the studies were largely school-based educational programmes and represented a somewhat homogenous evidence base. However, through our analysis, we were able to describe in detail the great variations in how these studies had been designed to achieve a common goal of reducing AAS use. We demonstrated the variety in intervention content, with 18 components included in different combinations across 14 interventions. The diversity in intervention design highlighted to us that there was little consensus as to how interventions in this field should attempt to reduce AAS use. This illustrates that examining the design of interventions can offer new insights into an evidence base that may not be immediately apparent without this level of scrutiny. Reflecting on this experience, we felt it helped to question our assumptions about what we knew about these interventions and an unintended benefit of scrutinising their design and content was that as reviewers we gained a greater understanding of the field.

As we were unable to undertake a pooled statistical analysis, we were only able to make cautious suggestions about intervention content that might be effective and worthy of consideration by others looking to develop or adapt their own similar intervention. Other researchers faced with similar limitations in data have done similar (Nyman et al., 2018). Where possible, however, reviewers can take this analysis a step further by exploring through statistical analyses the association between specific components of an intervention and its effectiveness. In addition to helping the reader understand the interventions, this can help identify with more confidence factors that are likely to lead to variation in the strength or direction of intervention effect. In a systematic review investigating the effectiveness of physical activity interventions delivered to inactive adults, the authors pooled effect sizes from 16 studies and coded the components utilised in these interventions (Howlett et al., 2019). They identified that collectively the interventions appeared effective at changing physical activity outcomes and through meta-regression identified
that effectiveness was associated with specific content including incorporating action planning, instruction on how to perform the behaviour, prompts/ cues, behaviour practice/ rehearsal, graded tasks, and self-reward. This supports reviewers to make recommendations not just about whether a broad intervention approach is likely to be effective, but what specific components are important. This is critical if reviewers want to make sense of large evidence bases that contain studies that have tested interventions consisting of many different components and combinations of them. A recent overview provides excellent discussion of the application of meta-regression and other statistical approaches to analyse intervention components in a meta-analysis (Petropoulou et al., 2021).

**Analysing use of theory**

In the context of interventions, theories help to explain why an intervention may lead to the intended outcomes, such as enhanced performance, motivation, self-efficacy or a change in behaviour. Reviewers may wish to explore whether interventions have a theoretical basis, the extent that they are used in the development of the intervention, and what theories are associated with effective or ineffective interventions. Similar to exploring intervention components, this scrutiny helps reviewers and the reader to have greater understanding about how interventions were designed to work and why they were associated with changes in outcomes. Given the plethora of named theories and the theoretical constructs they are based upon that those developing sport and exercise psychology interventions can utilise, it can be challenging to understand what theories to use (Michie et al., 2018) so exploring this in a systematic review provides valuable information.

Analysing the use of theory is therefore a logical step for reviewers, but one that can be challenging. It is often stated that interventions based on theory are likely to be more effective (Craig et al., 2008; Glanz & Bishop, 2010; Prestwich et al., 2015), but it can be difficult to demonstrate this association. The reasons for this vary, but commonly include that the theory may not have been fully applied in intervention development (Prestwich et al., 2015) or that the theory chosen may be inappropriate or simply not effective for bringing about the intended changes (Moore & Evans, 2017). A well-designed intervention with clear and appropriate application of a theory or theories may still fail to bring about the desired outcomes if the theory itself is flawed or limited. For example, a commonly applied theory to understand health-related behaviours has been the theory of planned behaviour (Ajzen, 1985, 1991), the strengths and weaknesses of which are highlighted in a recent and thorough overview of its application within sport psychology (Conner, 2020). The theory has been subject to much critique and has faced calls for it to be retired from use (Sniehotta et al., 2014) and reviewers are well placed to challenge or reinforce the use of specific theories such as this when they can identify whether their application is associated with effectiveness.

There is likely to be variation not just in whether any theory has been used to develop interventions, but how it is used with theory only applied to a limited extent in the development of many interventions (Prestwich et al., 2015). Reviewers hoping to assess the role of theory and its association with intervention effectiveness may therefore need to go further than simply identifying whether study authors allude to a theory or theories, but to examine how theory has been used (Michie & Prestwich, 2010). For example, has theory been used loosely by the study authors or has it informed the intervention design and have key theoretical constructs been tested in the evaluation?
Challenges in analysing intervention design

A significant challenge for reviewers hoping to analyse interventions in this manner and to identify effective intervention design is the quality of reporting they may encounter in included studies. All reviewers will be familiar with the challenges involved with making sense of journal articles that vary greatly in quality and reporting. Those hoping to scrutinise the minutia of intervention design and content require substantial detail in how study authors report their intervention. Reviewers can only, after all, analyse what is reported in publicly available materials. Unsurprisingly perhaps, they may find that many studies lack both detail and specificity about how an intervention was expected to work. As part of an interesting and critical reflection on their own work, Ely and colleagues describe that a lack of detail and too much ambiguity in their reporting of sport psychology studies limits uptake by others looking to adopt these interventions (Ely et al., 2020). Similarly, the authors of a systematic review looking at school-based physical activity highlighted that superficial description of interventions was a substantial barrier to their ability to make conclusions regarding the importance of specific components of these interventions (Hynynen et al., 2016).

It is not unusual to find that descriptions of interventions are inconsistent and reduced to short overviews that offer little more than an outline of what was implemented, or the overall focus of activities. More lengthy descriptions are often frustrating to try and make sense of when they lack specificity. For example, reviewers might be offered a thorough description of the broad approach and delivery of an educational intervention such as its goals, setting and provider, and the number of sessions over which it was delivered. However, without an account of the content of these sessions, they are left largely none the wiser about what participants experienced or how the intervention was expected to work and may not be able to identify with confidence the specific components present. Where authors use a range of often unclear or ambiguous terms to describe intervention content, it can be particularly challenging for reviewers to interpret these descriptions and there is a risk of miscoding components. This emphasises the need for reviewers to be reflexive about the studies they analyse and the limitations of the methods that they are using.

Such problems with reporting are particularly the case for older studies produced before tools such as the Template for Intervention Description and Replication (TIDieR) checklist (Hoffmann et al., 2014) encouraged detail in the reporting of interventions, and before journals could give authors the opportunity to publish supplementary materials online in addition to their manuscripts. Issues of superficial reporting and lack of detail will be nothing new to those undertaking systematic reviews and while it is important to recognise the limitations of the analysis it should not be seen as a barrier to scrutinising interventions in this way. Instead, the weaknesses of the evidence should be highlighted and used to create research recommendations to improve the reporting of interventions so that they can be better understood.

Complexity in environmental context

Context can be defined as the wider system that an intervention is implemented in, including geographical, cultural, social, organisational and political environments (Petticrew et al., 2013). Understanding how such factors influence how the intervention
works and its effectiveness clearly help us to understand these interventions and make recommendations about them. While this focus on context and implementation represents a departure from the focus on ‘does it work?’ in what remains arguably the dominant approach to analysing data in a systematic review - meta-analysis - it is an important step to take for reviews that aim to understand intervention effects and complexity.

As with all issues of complexity, reviewers will be interested in explaining how the variation in the environment that studies are set in influences the intervention and its effectiveness. For example, there have been many interventions designed to reduce sedentary behaviour in office-based workplaces in recognition of how sedentary working practices increase the risk of a range of negative health outcomes (Sui et al., 2019). To understand their impact reviewers can of course look at the content of what was delivered as part of these interventions and identify whether intervention components or mode of delivery, for example, appear to influence outcomes. However, beyond the intervention itself many other factors across the socioecological spectrum are likely to influence sedentary behaviours at work (Morris et al., 2018) and factors such as workplace culture and norms are significant influences (Owen et al., 2011). The implication is that the same intervention, such as the introduction of ‘standing desks’ or reminders to go for a walk at specific time points throughout the day, may have different impacts in different workplaces. More widely, reviewers can think about how community and societal-level factors in the bigger system that workplaces are set in can further influence the impact of such an intervention.

If the effects of an intervention are dependent on its context, this has implications for those hoping to replicate and adapt it in their own context. Exploring this complexity is not straightforward in systematic reviews and has not traditionally been the focus of review methods. Recently however, there has been greater interest in incorporating complexity into reviewing, as illustrated by a series of articles on the topic on behalf of the World Health Organisation (Higgins et al., 2019; Petticrew et al., 2019) that discuss the implications on review methods from taking a complexity or systems perspective. This engagement with complexity features most prominently in the health sciences, where reviewers commonly synthesise evidence from publicly delivered ‘real world’ interventions.

Recognising the complex systems that interventions exist in direct attention to considering how the characteristics of this system that may explain variation in how interventions work. This recognises that how an intervention is implemented and the variations in the environment or context that the intervention is implemented in contributes to complexity (Anderson et al., 2013) and that to understand interventions requires us to account for these differences sources of complexity (Petticrew et al., 2013). Analysing context complexity in systematic reviews is a relatively novel concept and as the emphasis grows on understanding complexity amongst leading groups such as Cochrane, it is likely that methods will continue to be developed. Reviewers asking questions about context and implementation will need to adapt their analysis approach to incorporate the types of evidence that will support them to answer such questions and to understand how variation in intervention effects might be explained by the wider system. A synthesis of findings generated through multiple analysis methods may be needed in systematic reviews that seek to explore complexity from multiple types of evidence, commonly discussed as mixed-methods synthesis (Petticrew et al., 2013). For example, a meta-analysis can be used alongside an analysis of complexity (Greenhalgh et al., 2011).
Analysing complexity in intervention implementation

Systematic reviews can include a focus on exploring implementation factors to provide greater insight towards explaining differences between interventions and their impacts. Reviewers seek to identify factors such as who is delivering an intervention, the setting, participants, fidelity (whether it was delivered as it was intended to be) and exposure (the extent to which the intended population received the intervention) because such factors can have an important role in mediating the effect on outcomes. For example, how does the impact of an intervention that aims to reduce burnout vary when it is delivered to university students or to elite performers? Or when it is delivered by a sports psychologist or the athletes’ coach? Can any variation in effectiveness be explained by the extent that the athletes engaged with the intervention, such as how many sessions of a programme they attended or how acceptable they found the intervention to be? It is not unusual for reviews to tackle these types of questions, but it is not currently done consistently and it is not without challenges.

While reporting of implementation factors in primary studies may be becoming more common it is still typically very poor. Reviewers often lack the information to determine factors such as the extent that interventions were delivered as they were intended, the level of exposure to the intervention amongst participants, or whether participants engaged with the intervention. In our review of AAS prevention interventions (Bates et al., 2019), we aimed to examine intervention fidelity but were unable to determine whether the interventions were delivered as intended or how participants experienced them, as relevant outcomes were only reported in one study. This is not an uncommon problem. For example, in a review examining psychological interventions to reduce injuries, the authors hoped to examine athletes’ compliance with interventions as a key outcome, but were limited by very infrequent reporting of this outcome (Gledhill et al., 2018). Qualitative explorations or mixed-methods studies of interventions are often more likely to provide evidence for reviewers who wish to analyse implementation factors, but are less commonly available than studies that report intervention effects.

Part 2: Guidelines for best practice

In this section, approaches to analyse complexity in a systematic review of interventions are outlined, based on four key recommendations that draw on key methods and tools, summarised in Table 1. While not a comprehensive list of all methods that can support analysis of complexity, these examples are excellent ways that reviewers can engage with complexity in design, implementation and environmental context as appropriate. Systematic reviews should always be guided by the need to provide evidence that their intended audience will find useful, interesting, and informative; and keeping the purpose of the review and research questions in mind is critical when identifying analysis methods.

Complexity in design

(i) Extract detailed information about interventions

In all systematic reviews of interventions, it is important to provide the reader with sufficient information to understand, develop or adapt the included interventions
The aim is not to overwhelm with data, but to provide what the reader will need to fully understand and, where appropriate, replicate or adapt what was introduced. One approach to guide the extraction of data about complex interventions in a systematic and thorough manner is to use the TIDieR checklist, which was designed to support the reporting of intervention design, materials, delivery and implementation (Hoffmann et al., 2014). Shepperd and colleagues (2009) recommended that for review analysis to be useful, detail is required on intervention content, fidelity and sustainability, and on the context and conditions that interventions are introduced in (Shepperd et al., 2009). Importantly, this analysis must be accurate and consistent across the studies included in the review.

Table 1. Key methods and tools to analyse complexity in intervention design, context, and implementation.

| Method, concept or tool | Use for analysing complexity in a systematic review | Examples of application within SEP literature |
|------------------------|-------------------------------------------------|---------------------------------------------|
| Intervention Complexity Assessment Tools for Systematic Reviews (iCAT_SR). Tool based on ten dimensions of intervention complexity (Lewin et al., 2017). | Guiding systematic and detailed assessment and categorisation of levels of complexity to provide an overall understanding of intervention complexity, which can be represented visually. | Blackburn et al. (2020); Jones et al. (2019) |
| Behaviour Change Taxonomy (Michie et al., 2013). Extensive taxonomy of 93 distinct behaviour change techniques (BCTs) that represent the active content of interventions. | Supporting consistent identification and comparison of the components or ‘active ingredients’ that interventions are built on. BCTs can be used in subgroup analysis to identify effective components, such as through meta-regression. | Howlett et al. (2019); Knittle et al. (2018) |
| Theory Coding Scheme (Michie & Prestwich). 19-item tool to assess the extent which interventions are theory-based. Meta-regression (Baker et al., 2009) to explore the relative effectiveness of different intervention components. | Supporting consistent identification and comparison of the use of theory in intervention design and implementation. Identifying the components of interventions (such as BCTs) that appear effective. Other advanced statistical methods such as network meta-analysis can be used to further explore complexity in intervention design. | Allan et al. (2018); Bentley et al. (2020) |
| Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. 12-item reporting standards to guide detailed reporting of interventions (Hoffmann et al., 2014) | Guiding consistent extraction of intervention details including procedures, setting, delivery, implementation, and process outcomes. | Bentley et al. (2020); Klempel et al. (2021) |
| Logic models exploring the underlying theories of interventions and system-level influences intervention outcomes (Petticrew et al., 2019) | Supporting understanding of how the wider system works, and how context influences intervention implementation and outcomes. Identifying outcomes to include in a review. Communicating complexity to readers. | Kahn et al. (2002); Motamed-Jahromi and Kaveh (2021) |
| Realist synthesis (Pawson & Bellamy, 2006) to explore the underlying causal mechanisms that explain why interventions work or do not work. | Supporting understand of how environmental context and intervention implementation factors influences the effectiveness of findings. Can be carried out alongside a review of effectiveness outcomes. | Leone and Pesce (2017); Pesce et al. (2021) |
| Measures of factors influencing how interventions are implemented or experienced. For example, dimensions of implementation (Cargo et al., 2018) including quantitative and qualitative measures. | Identifying implementation outcomes to help explain variation in effects. Typically included in reviews of qualitative or mixed-methods evidence, but can be included alongside a review of effectiveness outcomes. | Farrance et al. (2016); Naylor et al. (2015) |
When criteria in the TIDieR checklist are used to extract the detail of interventions in a review it can provide excellent detail and overview of the evidence base, as demonstrated in many recent review articles from the sport and exercise psychology literature (Bentley et al., 2020; Lim et al., 2019; Madden et al., 2020; Van Zyl et al., 2020). Using TIDieR or similar as part of the process of data extraction should not require any additional burden of resources upon reviewers and will help develop their own understanding of the evidence. Undertaking a thorough assessment of the interventions not only provides the readers and reviewers with important information to understand the interventions and highlights any limitations in study reporting, but also helps to examine complexity. The constructs in the TIDier checklist support us to start thinking about these important questions of what the intervention involved, how it was delivered, and where and how it was implemented.

An important function of systematic reviews is to highlight limitations and gaps in reporting, and to drive efforts to improve this. There are no easy solutions to overcome the often-inadequate level of intervention description in primary studies. However, reviewers may find more substantial descriptions of interventions published separately in protocols, guides, or as supplementary materials to the study. Reviewers can also consider contacting study authors and asking for missing information to gain greater clarity and detail in addition to published information. In our review of AAS prevention interventions we received responses for further details about the design of interventions from five of the six authors we contacted. However, these responses provided little further insight as authors were unable to, or choose not to, provide much information beyond that already published.

(ii) Use behavioural science tools to code intervention content and design

Reviewers hoping to analyse intervention design will benefit from engaging with tools designed to support the systematic and consistent coding of interventions. This will support their efforts to provide sufficient detail about interventions, but beyond this to be explicit about their design and explore what factors are associated with variations in effectiveness. One tool that systematic reviewers have increasingly used to identify the content of interventions is the Behaviour Change Technique Taxonomy (Michie et al., 2013). It was designed to be used to code and report the smallest components of interventions, sometimes referred to as their ‘active ingredients’, across different disciplines and fields. The taxonomy includes 93 ‘behaviour change techniques’ (BCTs) that represent these active ingredients and point to the potential for huge variety in intervention design based on combinations of these BCTs. Being specific about the content will help us understand how interventions were intended to work, and what components appear important. To use it requires skill and confidence in recognising and coding intervention components made more difficult by inconsistent use of terminology across primary studies. The authors of the BCT taxonomy offer a free and comprehensive training resource (Michie et al., 2021) that provides valuable support on how to identify and code BCTs.

The authors specify that a key advantage of identifying BCTs is that it supports reviewers to identify how intervention design is associated with effectiveness, and applying the BCT taxonomy to identify and analyse the importance of intervention components has been a popular approach recently in the field of sport and exercise psychology.
(Bentley et al., 2020; Dombrowski et al., 2012; Flannery et al., 2019; Gardner et al., 2011). As more detail is sought for and included about interventions, it becomes important to find concise and interesting ways to present this data to avoid relying on lengthy summary tables that can be challenging to digest. For example, when reporting somewhat dense data such as BCTs across large numbers of studies, reviewers can present this in easy-to-understand summary figures that show not only which BCTs featured in which studies, but give a sense of how common each BCT is in the evidence base and whether they seem promising or otherwise (Nyman et al., 2018).

A second tool that can be used to explore intervention design is the Theory Coding Scheme (Michie & Prestwich, 2010), which was designed to help examine the extent to which interventions are theoretically based in their design, implementation and evaluation. When applied in a systematic review it provides a systematic approach to assess the role of theory consistently across different studies. For example, it has been used by reviewers to make recommendations to increase the explicit use of theory in the design of coach development programmes, and for the rigorous evaluation of these theoretical constructs (Allan et al., 2018). The authors of the Theory Coding Scheme applied it themselves in a review of 140 physical activity and healthy eating interventions and identified that not only was theory rarely used extensively as part of intervention development, but through meta-regression demonstrated that there was little association between the use of theory and intervention effectiveness (Prestwich et al., 2014).

**Complexity in environmental context**

(iii) *Engage with logic models*

Logic models can be an excellent tool to plan for and explore complexity in systematic review analysis, and to communicate complexity to the reader. They map out the important concepts in a systematic review and propose links between the intervention and outcomes to demonstrate how interventions are proposed to work (Baxter et al., 2014; Kneale et al., 2015). Anderson and colleagues (2011) discuss how logic models can therefore provide an evidence-based ‘analytical map’ that help reviewers to identify the outcomes they include to judge effectiveness and identify contextual factors that might contribute to variation in how interventions are experienced and lead to changes in outcomes. When thinking of possible outcomes, models can be built by working backwards from potential distal outcomes that could change after the intervention along a causal path to identify intermediate outcomes and the measurable outputs of the interventions (Kneale et al., 2015). In a well-cited review of interventions to increase physical activity (Kahn et al., 2002), the authors present a good example of a simple logic model that show how these interventions may lead to distal health outcomes via the intermediate physiological outcomes associated with changes in physical activity. Logic models can further support analysis by providing structure to this part of the review process (Rohwer et al., 2017). When the model indicates that different groups may experience variations in how an intervention effects them, this provides support and justification for sub-group analyses (Anderson et al., 2011; Rohwer et al., 2017) and they can guide narrative or thematic
synthesis by identifying important and relevant areas to focus most attention on (Kneale et al., 2015).

Logic models may be particularly useful to help understand how interventions in complex environments might lead to long-term change. The process of building the model will often be of great help to reviewers to understand complexity in the intervention or its implementation. For those hoping to use the evidence provided in a systematic review, they can help clarify how interventions are intended to work through graphical representations of the key intervention components and the causal pathways between these components and outcomes (Baxter et al., 2014; Rehfuess et al., 2018; Rohwer et al., 2017). A well-presented logic model can provide transparency about the assumptions reviewers make on how the intervention design or context influences outcomes, which can be challenging to convey in text. This improves the accessibility of evidence presented (Rohwer et al., 2017).

While they have a useful function in supporting analysis in systematic reviews, logic models have utility beyond supporting this such as in formulating review questions, identifying inclusion criteria and guiding search strategies (Anderson et al., 2011). While the benefits of engaging with logic models are clear, it is important to recognise that when reviewers draw on one it will have a large influence on how the review progresses and developing a model can be a substantial undertaking in itself (Rohwer et al., 2017). Reviewers may find logic models already available in the literature suitable for their review, or they can build new ones through their understanding of the interventions and drawing on evidence from a scoping review or rapid synthesis of relevant evidence. The articles cited in this section provide useful advice and guidance around developing and utilising logic models in a review.

**Complexity in implementation**

(iv) **Integrate implementation outcomes with analysis of intervention effectiveness**

Guidance from the Cochrane Qualitative and Implementation Methods Group provides examples of dimensions of intervention implementation that reviewers can consider such as reach, fidelity, adaptation and engagement; with both quantitative and qualitative indicators that can guide identification of outcomes (Cargo et al., 2018). Studies of interventions increasingly provide some measures of implementation alongside measures of effectiveness, and it is not only reviews that specifically seek data from qualitative or mixed-methods studies that can incorporate these outcomes. However, implementation outcomes are most likely to be drawn from studies of interventions that include a qualitative element and reviewers interested in complexity in implementation should consider how to integrate evidence from qualitative or mixed-methods studies alongside evidence of intervention effectiveness. The Cochrane Qualitative and Implementation Methods Group provide guidance to support this (Harden et al., 2018).

The work of the RAMESES group (Greenhalgh et al., 2011; Wong et al., 2014) is highly relevant to consider when seeking to explore complexity of intervention implementation and environmental context. As Greenhalgh and colleagues (2011) outline, a traditional Cochrane-type review of effectiveness can be undertaken alongside (for example) a
realist review approach that will together provide an understanding of ‘what works’ and ‘why did it work’. The group focuses on theory-driven qualitative and mixed-method reviews that seek to identify and understand the factors that influence how interventions work. Pawson’s realist synthesis (Pawson & Bellamy, 2006; Pawson & Tilley, 1997) is of particular relevance to those seeking to understand the importance of the context that interventions are introduced in, with its oft-repeated realist-inspired questions of what works, how, for whom, in what circumstances? When considering interventions to increase physical activity amongst young children based on a range of approaches such as mentoring, goal setting and incentivisation, Hnatiuk and colleagues used meta-analysis to identify changes in physical activity, and subgroup analyses revealed the importance of setting and parental components (Hnatiuk et al., 2019). They included a realist synthesis component to explore implementation and contextual factors, which offered insight into how to deliver the content of intervention and to overcome barriers to implementation, and evidence around the ideal frequency and duration of intervention sessions. The authors used the findings of the realist synthesis to explain why interventions were, or were not, effective at changing children’s physical activity, offering an excellent example of how analysing complexity in implementation and context can improve understanding and usefulness of effectiveness findings.

Whether or not the intention is to undertake some variation of realist or mixed-methods review, considering realist approaches provides invaluable insight for reviewers who are interested in the role of variation in context on interventions. For many reviewers, the important implications may be in the study designs that they include in their synthesis and the types of outcomes that they seek to identify. Including outcomes that help to understand implementation and complexity and exploring their association with intervention effectiveness will provide important insight for reviewers hoping to understand the context that interventions are delivered in.

**Conclusion**

As understanding and interest in the complexity of the design, environmental context, and implementation of interventions continues to grow, systematic reviewers need to respond in how they analyse sport and exercise psychology interventions. When reviewers engage with complexity they can continue to synthesise and make recommendations about the issues that matter to decision-makers and those who develop and adapt interventions. Systematic review methods continue to be developed that will support reviewers to understand interventions and explain their effectiveness through analysis of this complexity. Incorporating complexity in systematic reviews requires reviewers to ask different questions and engage with methods that support analysis that can answer these questions, including qualitative or mixed methods data, as well as methods that examine the variation in intervention design. While complexity can be challenging to incorporate into an analysis of interventions, it is how reviewers can strive to answer important questions of how and why interventions work, and what factors predict variation in effectiveness.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).
References

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Eds.), *Action control* (pp. 11–39). Springer.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Allan, V., Vierimaa, M., Gainforth, H. L., & Côté, J. (2018). The use of behaviour change theories and techniques in research-informed coach development programmes: A systematic review. *International Review of Sport and Exercise Psychology*, 11(1), 47–69. https://doi.org/10.1080/1750984X.2017.1286514

Anderson, L. M., Oliver, S. R., Michie, S., Rehfuess, E., Noyes, J., & Shemilt, I. (2013). Investigating complexity in systematic reviews of interventions by using a spectrum of methods. *Journal of Clinical Epidemiology*, 66(11), 1223–1229. https://doi.org/10.1016/j.jclinepi.2013.06.014

Anderson, L. M., Petticrew, M., Rehfuess, E., Armstrong, R., Ueffing, E., Baker, P., & Tugwell, P. (2011). Using logic models to capture complexity in systematic reviews. *Research Synthesis Methods*, 2(1), 33–42. https://doi.org/10.1002/jrsm.32

Baker, W. L., White, C. M., Cappelleri, J. C., Kluger, J., & Coleman, C. I. (2009). Understanding heterogeneity in meta-analysis: The role of meta-regression. *International Journal of Clinical Practice*, 63(10), 1426–1434. https://doi.org/10.1111/j.1742-1241.2009.02168.x

Bates, G., Begley, E., Tod, D., Jones, L., Leavey, C., & McVeigh, J. (2019). A systematic review investigating the behaviour change strategies in interventions to prevent misuse of anabolic steroids. *Journal of Health Psychology*, 24(11), 1595–1612. https://doi.org/10.1177/1359105317737607

Baxter, S. K., Blank, L., Woods, H. B., Payne, N., Rimmer, M., & Goyder, E. (2014). Using logic model methods in systematic review synthesis: Describing complex pathways in referral management interventions. *BMC Medical Research Methodology*, 14(1), 62. https://doi.org/10.1186/1471-2288-14-62

Bentley, M. R. N., Mitchell, N., & Backhouse, S. H. (2020). Sports nutrition interventions: A systematic review of behavioural strategies used to promote dietary behaviour change in athletes. *Appetite*, 150, 104645. https://doi.org/10.1016/j.appet.2020.104645

Blackburn, N. E., Wilson, J. J., McMullan, I. I., Caserotti, P., Giné-Garriga, M., Wirth, K., & Tully, M. A. (2020). The effectiveness and complexity of interventions targeting sedentary behaviour across the lifespan: A systematic review and meta-analysis. *International Journal of Behavioral Nutrition and Physical Activity*, 17(1), 53. https://doi.org/10.1186/s12966-020-00957-0

Campbell, M., Katikireddi, S. V., Sowden, A., & Thomson, H. (2019). Lack of transparency in reporting narrative synthesis of quantitative data: A methodological assessment of systematic reviews. *Journal of Clinical Epidemiology*, 105, 1–9. https://doi.org/10.1016/j.jclinepi.2018.08.019

Campbell, M., McKenzie, J. E., Sowden, A., Katikireddi, S. V., Brennan, S. E., Ellis, S., & Thomson, H. (2020). Synthesis without meta-analysis (SWIM) in systematic reviews: Reporting guideline. *BMJ*, 368, l6890. https://doi.org/10.1136/bmj.l6890

Cargo, M., Harris, J., Pantoja, T., Booth, A., Harden, A., Hannes, K., & Noyes, J. (2018). Cochrane Qualitative and Implementation Methods Group guidance series—paper 4: Methods for assessing evidence on intervention implementation. *Journal of Clinical Epidemiology*, 97, 59–69. https://doi.org/10.1016/j.jclinepi.2017.11.028

Chandler, J., Cumpston, M., Thomas, J., Higgins, J.P.T., Deeks, J. J., & Clarke, M.J. (2021). Chapter 1: Introduction. In J.P.T. Higgins, J. Thomas, J Chandler, M. Cumpston, T. Li, M.J. Page, & V.A. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions Version 6.2*. Cochrane.

Conner, M. (2020). Theory of planned behavior. In G. Tenenbaum & R.C. Eklund (Eds.), *Handbook of Sport Psychology* (pp. 1–18)

Craig, P., Dieppe, P., MacIntyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*, 337, a1655. https://doi.org/10.1136/bmj.a1655
Deeks, J. J., Higgins, J. P. T., Altman, D. G. (2019). Analysing data and undertaking meta-analyses. In J.P. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M.J. Page, & V.A. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions*. Cochrane.

Dombrowski, S. U., Sniehotta, F. F., Avenell, A., Johnston, M., MacLennan, G., & Araújo-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(1), 7–32. https://doi.org/10.1080/17437199.2010.513298

Eather, N., Babic, M., Riley, N., Harris, N., Jung, M., Jeffs, M., & Lubans, D. R. (2020). Integrating high-intensity interval training into the workplace: The work-HIIT pilot RCT. *Scandinavian Journal of Medicine & Science in Sports*, 30(12), 2445–2455. https://doi.org/10.1111/sms.13811

Eckerstorfer, L. V., Tanzer, N. K., Vogrincic-Haselbacher, C., Kedia, G., Brohmer, H., Dinslaken, I., & Corcoran, K. (2018). Key elements of mHealth interventions to successfully increase physical activity: Meta-regression. *JMIR mHealth and uHealth*, 6(11), 2291–5222. https://doi.org/10.2196/10076

Ely, F. O., Jenny, O., & Munroe-Chandler, K. J. (2020). How intervention research designs may broaden the research-to-practice gap in sport psychology. *Journal of Sport Psychology in Action*, 12(2), 101–113. https://doi.org/10.1080/21520704.2020.1798573

Farrance, C., Tsolfiou, F., & Clark, C. (2016). Adherence to community based group exercise interventions for older people: A mixed-methods systematic review. *Preventive Medicine*, 87, 155–166. https://doi.org/10.1016/j.ypmed.2016.02.037

Flannery, C., Fredrix, M., Olander, E. K., McAuliffe, F. M., Byrne, M., & Kearney, P. M. (2019). Effectiveness of physical activity interventions for overweight and obesity during pregnancy: A systematic review of the content of behaviour change interventions. *International Journal of Behavioral Nutrition and Physical Activity*, 16(1), 97. https://doi.org/10.1186/s12966-019-0859-5

Franco, M. R., Tong, A., Howard, K., Sherrington, C., Ferreira, P. H., Pinto, R. Z., & Ferreira, M. L. (2015). Older people’s perspectives on participation in physical activity: A systematic review and thematic synthesis of qualitative literature. *British Journal of Sports Medicine*, 49(19), 1268. https://doi.org/10.1136/bjsports-2014-094015

Gardner, B., Wardle, J., Poston, L., & Croker, H. (2011). Changing diet and physical activity to reduce gestational weight gain: A meta-analysis. *Obesity Reviews*, 12(7), e602–e620. https://doi.org/10.1111/j.1467-789X.2011.00884.x

Glanz, K., & Bishop, D. B. (2010). The role of behavioral science theory in development and implementation of public health interventions. *Annual Review of Public Health*, 31(1), 399–418. https://doi.org/10.1146/annurev.publhealth.012809.103604

Gledhill, A., Forsdyke, D., & Murray, E. (2018). Psychological interventions used to reduce sports injuries: A systematic review of real-world effectiveness. *British Journal of Sports Medicine*, 52(15), 967. https://doi.org/10.1136/bjsports-2017-097694

Greenhalgh, T., Wong, G., Westhorp, G., & Pawson, R. (2011). Protocol - realist and meta-narrative evidence synthesis: Evolving Standards (RAMESES). *BMC Medical Research Methodology*, 11(1), 115. https://doi.org/10.1186/1471-2288-11-115

Harden, A., Thomas, J., Cargo, M., Harris, J., Pantoja, T., Flemming, K., & Noyes, J. (2018). Cochrane Qualitative and Implementation Methods Group guidance series-paper 5: Methods for integrating qualitative and implementation evidence within intervention effectiveness reviews. *Journal of Clinical Epidemiology*, 97, 70–78. https://doi.org/10.1016/j.jclinepi.2017.11.029

Higgins, J. P. T., López-López, J. A., Becker, B. J., Davies, S. R., Dawson, S., Grimshaw, J. M., & Caldwell, D. M. (2019a). Synthesising quantitative evidence in systematic reviews of complex health interventions. *BMJ Global Health*, 4(Suppl 1), e000858. https://doi.org/10.1136/bmjgh-2018-000858

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2019b). *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons.

Hnatiuk, J. A., Brown, H. E., Downing, K. L., Hinkley, T., Salmon, J., & Hesketh, K. D. (2019). Interventions to increase physical activity in children 0–5 years old: A systematic review, meta-analysis and realist synthesis. *Obesity Reviews*, 20(1), 75–87. https://doi.org/10.1111/obr.12763
Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., & Michie, S. (2014). Better reporting of interventions: Template for intervention description and replication (TIDieR) checklist and guide. *BMJ*, 348, g1687. https://doi.org/10.1136/bmj.g1687

Hoffmann, T. C., Oxman, A. D., Ioannidis, J. P. A., Moher, D., Lasserson, T. J., Tovey, D. I., & Glasziou, P. (2017). Enhancing the usability of systematic reviews by improving the consideration and description of interventions. *BMJ*, 358, j2998. https://doi.org/10.1136/bmj.j2998

Howlett, N., Trivedi, D., Troop, N. A., & Chater, A. M. (2019). Are physical activity interventions for healthy inactive adults effective in promoting behavior change and maintenance, and which behavioral change techniques are effective? A systematic review and meta-analysis. *Translational Behavioral Medicine*, 9(1), 147–157. https://doi.org/10.1093/tbm/iby010

Hut, M., Glass, C. R., Degnan, K. A., & Minkler, T. O. (2021). The effectiveness of e-health interventions promoting physical activity in older people: A systematic review. *International Journal of Environmental Research and Public Health*, 18(4), 1902. https://doi.org/10.3390/ijerph18041902

Hynynen, S. T., van Stralen, M. M., Sniehotta, F. F., Araújo-Soares, V., Hardeman, W., Chinapaw, M. J. M., & Hankonen, N. (2016). A systematic review of school-based interventions targeting physical activity and sedentary behaviour among older adolescents. *International Review of Sport and Exercise Psychology*, 9(1), 22–44. https://doi.org/10.1080/17437199.2015.1081706

Jones, R. A., Blackburn, N. E., Woods, C., Byrne, M., van Nassau, F., & Tully, M. A. (2019). Interventions promoting active transport to school in children: A systematic review and meta-analysis. *Preventive Medicine*, 123, 232–241. https://doi.org/10.1016/j.ypmed.2019.03.030

Kahn, E. B., Ramsey, L. T., Brownson, R. C., Heath, G. W., Howze, E. H., Powell, K. E., & Corso, P. (2002). The effectiveness of interventions to increase physical activity: A systematic review. *American Journal of Preventive Medicine*, 22(4), 73–107. https://doi.org/10.1016/S0749-3797(02)00434-8

Kok, G., Gottlieb, N. H., Peters, G.-J. Y., Mullen, P. D., Parcel, G. S., Ruiter, R. A. C., & Bartholomew, L. K. (2016). A taxonomy of behaviour change methods: An intervention mapping approach. *Health Psychology Review*, 10(3), 297–312. https://doi.org/10.1080/17437199.2015.1077155

Kriemler, S., Meyer, U., Martin, E., van Sluijs, E. M. F., Andersen, L. B., & Martin, B. W. (2011). Efficacy of chair-based exercise on physical function in older adults: A systematic review and meta-analysis. *International Journal of Environmental Research and Public Health*, 8(4), 1902. https://doi.org/10.3390/ijerph18041902

Kneale, D., Thomas, J., & Harris, K. (2015). Developing and optimising the use of logic models in systematic reviews: Exploring practice and good practice in the use of programme theory in reviews. *PLoS One*, 17(10), e0142187. https://doi.org/10.1371/journal.pone.0142187

Knittle, K., Nurmi, J., Crutzen, R., Hankonen, N., Beattie, M., & Dombrowski, S. U. (2018). How can interventions increase motivation for physical activity? A systematic review and meta-analysis. *Health Psychology Review*, 12(3), 211–230. https://doi.org/10.1080/17437199.2018.1435299

Kok, G., Gottlieb, N. H., Peters, G.-J. Y., Mullen, P. D., Parcel, G. S., Ruiter, R. A. C., & Bartholomew, L. K. (2016). A taxonomy of behaviour change methods: An intervention mapping approach. *Health Psychology Review*, 10(3), 297–312. https://doi.org/10.1080/17437199.2015.1077155

Krippner, S., Tenopir, C., & West, M. R. (2010). The effectiveness of e-health interventions promoting physical activity in older people: A systematic review and meta-analysis. *European Review of Aging and Physical Activity*, 7(1), 1–17. https://doi.org/10.1186/1186/s11556-019-0235-0

Kwan, R. Y. C., Salihu, D., Lee, P. H., Tse, M., Cheung, D. S. K., Roopsawang, I., & Choi, K. S. (2020). The effect of e-health interventions promoting physical activity in older people: A systematic review and meta-analysis. *European Review of Aging and Physical Activity*, 17(1), 1–17. https://doi.org/10.1186/s11556-019-0235-0

Kwan, R. Y. C., Salihu, D., Lee, P. H., Tse, M., Cheung, D. S. K., Roopsawang, I., & Choi, K. S. (2020). The effect of e-health interventions promoting physical activity in older people: A systematic review and meta-analysis. *European Review of Aging and Physical Activity*, 17(1), 1–17. https://doi.org/10.1186/s11556-019-0235-0

Kwan, R. Y. C., Salihu, D., Lee, P. H., Tse, M., Cheung, D. S. K., Roopsawang, I., & Choi, K. S. (2020). The effect of e-health interventions promoting physical activity in older people: A systematic review and meta-analysis. *European Review of Aging and Physical Activity*, 17(1), 1–17. https://doi.org/10.1186/s11556-019-0235-0

Kwan, R. Y. C., Salihu, D., Lee, P. H., Tse, M., Cheung, D. S. K., Roopsawang, I., & Choi, K. S. (2020). The effect of e-health interventions promoting physical activity in older people: A systematic review and meta-analysis. *European Review of Aging and Physical Activity*, 17(1), 1–17. https://doi.org/10.1186/s11556-019-0235-0
Lim, S., Liang, X., Hill, B., Teede, H., Moran, L. J., & O’Reilly, S. (2019). A systematic review and meta-analysis of intervention characteristics in postpartum weight management using the TIDieR framework: A summary of evidence to inform implementation. *Obesity Reviews, 20*(7), 1045–1056. https://doi.org/10.1111/obr.12846

Madden, S. K., Cordon, E. L., Bailey, C., Skouteris, H., Ahuja, K., Hills, A. P., & Hill, B. (2020). The effect of workplace lifestyle programmes on diet, physical activity, and weight-related outcomes for working women: A systematic review using the TIDieR checklist. *Obesity Reviews, 21*(10), e13027. https://doi.org/10.1111/obr.13027

Michie, S., Atkins, L., & West, R. (2014). *The behaviour change wheel. A guide to designing interventions*. Silverback Publishing.

Michie, S., Carey, R. N., Johnston, M., Rothman, A. J., de Bruin, M., Kelly, M. P., & Connell, L. E. (2018). From theory-inspired to theory-based interventions: A protocol for developing and testing a methodology for linking behaviour change techniques to theoretical mechanisms of action. *Annals of Behavioral Medicine, 52*(6), 501–512. https://doi.org/10.1007/s12160-016-9816-6

Michie, S., Johnston, M., Francis, J., Abraham, C., Hardeman, W., & Wood, C. (2021). *BCTTv1 Online Training*. Retrieved February 2, 2021, from www.bct-taxonomy.com

Michie, S., & Prestwich, A. (2010). Are interventions theory-based? Development of a theory coding scheme. *Health Psychology, 29*(1), 1–8. https://doi.org/10.1037/a0016939

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., & Wood, C. E. (2013). The behaviour 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*(1), 81–95. https://doi.org/10.1007/s12160-013-9486-6

Minary, L., Trompette, J., Kivits, J., Cambon, L., Tarquinio, C., & Alla, F. (2019). Which design to evaluate complex interventions? Toward a methodological framework through a systematic review. *BMC Medical Research Methodology, 19*(1), 1–9. https://doi.org/10.1186/s12874-019-0736-6

Moore, G. F., & Evans, R. E. (2017). What theory, for whom and in which context? Reflections on the application of theory in the development and evaluation of complex population health interventions. *SSM – Population Health, 3*, 132–135. https://doi.org/10.1016/j.ssmph.2016.12.005

Morris, A., Murphy, R., Shepherd, S., & Graves, L. (2018). Multi-stakeholder perspectives of factors that influence contact centre call agents’ workplace physical activity and sedentary behaviour. *International Journal of Environmental Research and Public Health, 15*(7), 1484. https://doi.org/10.3390/ijerph15071484

Motamed-Jahromi, M., & Kaveh, M. H. (2021). Effective interventions on improving elderly’s independence in activity of daily living: A systematic review and logic model. *Frontiers in Public Health, 8*, 516151. https://doi.org/10.3389/fpubh.2020.516151

Movsisyan, A., Arnold, L., Evans, R., Hallingberg, B., Moore, G., O’Cathain, A., & Rehfuess, E. (2019). Adapting evidence-informed complex population health interventions for new contexts: A systematic review of guidance. *Implementation Science, 14*(1), 105. https://doi.org/10.1186/s13012-019-0956-5

Munn, Z., Tufanaru, C., & Aromataris, E. (2014). JBI’s systematic reviews: Data extraction and synthesis. *AJN The American Journal of Nursing, 114*(7), 49–54. https://doi.org/10.1097/01.NAJ.0000451683.66447.89

Naylor, P.-J., Nettlefold, L., Race, D., Hoy, C., Ashe, M. C., Wharf Higgins, J., & McKay, H. A. (2015). Implementation of school based physical activity interventions: A systematic review. *Preventive Medicine, 72*, 95–115. https://doi.org/10.1016/j.ypmed.2014.12.034

Nilsen, P. (2015). Making sense of implementation theories, models and frameworks. *Implementation Science, 10*(1), 53. https://doi.org/10.1186/s13012-015-0242-0

Nyman, S. R., Adamczewska, N., & Howlett, N. (2018). Systematic review of behaviour change techniques to promote participation in physical activity among people with dementia. *British Journal of Health Psychology, 23*(1), 148–170. https://doi.org/10.1111/bjhp.12279

O’Cathain, A., Croft, L., Sworn, K., Duncan, E., Rousseau, N., Turner, K., & Hoddinott, P. (2019). Taxonomy of approaches to developing interventions to improve health: A systematic review.
methods overview. *Pilot and Feasibility Studies*, 5(1), 41. https://doi.org/10.1186/s40814-019-0425-6

Owen, N., Sugiyama, T., Eakin, E. E., Gardiner, P. A., Tremblay, M. S., & Sallis, J. F. (2011). Adults’ sedentary behavior: Determinants and interventions. *American Journal of Preventive Medicine*, 41(2), 189–196. https://doi.org/10.1016/j.amepre.2011.05.013

Pawson, R., & Bellamy, J. L. (2006). Realist synthesis: An explanatory focus for systematic review. In J. Popay (Ed.), *Moving Beyond Effectiveness in Evidence Synthesis* (pp. 83–94). London: National Institute for Health and Clinical Excellence.

Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. sage.

Pesce, C., Vazou, S., Benzing, V., Álvarez-Bueno, C., Anzeneder, S., Mavilidi, M. F., & Schmidt, M. (2021). Effects of chronic physical activity on cognition across the lifespan: A systematic meta-review of randomized controlled trials and realist synthesis of contextualized mechanisms. *International Review of Sport and Exercise Psychology*, 1–39. https://doi.org/10.1080/1750984X.2021.1929404

Petropoulou, M., Efthimiou, O., Rücker, G., Schwarzer, G., Furukawa, T. A., Pompoli, A., & Mavridis, D. (2021). A review of methods for addressing components of interventions in meta-analysis. *PLOS ONE*, 16(2), e0246631. https://doi.org/10.1371/journal.pone.0246631

Petticrew, M. (2011). When are complex interventions ‘complex’? When are simple interventions ‘simple’? *The European Journal of Public Health*, 21(4), 397–398. https://doi.org/10.1093/eurpub/ckr084

Petticrew, M., Knai, C., Thomas, J., Rehfues, E. A., Noyes, J., Gerhardus, A., & McGill, E. (2019). Implications of a complexity perspective for systematic reviews and guideline development in health decision making. *BMJ Global Health*, 4(Suppl 1), e000899. https://doi.org/10.1136/bmjgh-2018-000899

Petticrew, M., Rehfues, E., Noyes, J., Higgins, J. P. T., Mayhew, A., Pantoja, T., & Sowden, A. (2013). Synthesizing evidence on complex interventions: How meta-analytical, qualitative, and mixed-method approaches can contribute. *Journal of Clinical Epidemiology*, 66(11), 1230–1243. https://doi.org/10.1016/j.jclinepi.2013.06.005

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Araújo-Soares, L., Rodgers, M., & Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. A Product from the ESRC Methods Programme. Version 1: b92

Prestwich, A., Sniehotta, F. F., Whittington, C., Dombrowski, S. U., Rogers, L., & Michie, S. (2014). Does theory influence the effectiveness of health behavior interventions? Meta-analysis. *Health Psychology*, 33(5), 465. https://doi.org/10.1037/a0032853

Prestwich, A., Webb, T. L., & Conner, M. (2015). Using theory to develop and test interventions to promote changes in health behaviour: Evidence, issues, and recommendations. *Current Opinion in Psychology*, 5, 1–5. https://doi.org/10.1016/j.copsyc.2015.02.011

Rehfues, E. A., Booth, A., Brereton, L., Burns, J., Gerhardus, A., Mozygemba, K., & Rohwer, A. (2018). Towards a taxonomy of logic models in systematic reviews and health technology assessments: A priori, staged, and iterative approaches. *Research Synthesis Methods*, 9(1), 13–24. https://doi.org/10.1002/1jrs.1254

Richards, D. A., & Hallberg, I. R. (2015). *Complex interventions in health: An overview of research methods*. Routledge.

Rohwer, A., Pfadenhauer, L., Burns, J., Brereton, L., Gerhardus, A., Booth, A., & Rehfues, E. (2017). Logic models help make sense of complexity in systematic reviews and health technology assessments. *Journal of Clinical Epidemiology*, 83, 37–47. https://doi.org/10.1016/j.jclinepi.2016.06.012

Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn’t. *BMJ*, 312(7023), 71. https://doi.org/10.1136/bmj.312.7023.71

Shannon, S. A.-O., Brennan, D., Hanna, D., Younger, Z., Hassan, J., & Breslin, G. (2018). The effect of a school-based intervention on physical activity and well-being: A non-randomised controlled trial with children of low socio-economic status. *Sports Medicine - Open*, 4(1), 16. https://doi.org/10.1186/s40798-018-0129-0
Shepperd, S., Lewin, S., Straus, S., Clarke, M., Eccles, M. P., Fitzpatrick, R., & Sheikh, A. (2009). Can we systematically review studies that evaluate complex interventions? *PLoS Medicine, 6*(8), e1000086. https://doi.org/10.1371/journal.pmed.1000086

Sniehotta, F. F., Presseau, J., & Araújo-Soares, V. (2014). Time to retire the theory of planned behaviour. *Health Psychology Review, 8*(1), 1–7. https://doi.org/10.1080/17437199.2013.869710

Snijstveit, B., Oliver, S., & Vojtkova, M. (2012). Narrative approaches to systematic review and synthesis of evidence for international development policy and practice. *Journal of Development Effectiveness, 4*(3), 409–429. https://doi.org/10.1080/19439342.2012.710641

Sui, W., Smith, S. T., Fagan, M. J., Rollo, S., & Prapavessis, H. (2019). The effects of sedentary behaviour interventions on work-related productivity and performance outcomes in real and simulated office work: A systematic review. *Applied Ergonomics, 75*, 27–73. https://doi.org/10.1016/j.apergo.2018.09.002

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology, 8*(1), 45. https://doi.org/10.1186/1471-2288-8-45

Tod, D., & Eubank, M. R. (2017). Conducting a systematic review: Demystification for trainees in sport and exercise psychology. *Sport and Exercise Psychology Review, 13*(1).

Tricco, A. C., Tetzlaff, J., & Moher, D. (2011). The art and science of knowledge synthesis. *Journal of Clinical Epidemiology, 64*(1), 11–20. https://doi.org/10.1016/j.jclinepi.2009.11.007

Van Zyl, N., Andrews, L., Williamson, H., & Meyrick, J. (2020). The effectiveness of psychosocial interventions to support psychological well-being in post-operative bariatric patients: A systematic review of evidence. *Obesity Research & Clinical Practice, 14*(5), 404–420. https://doi.org/10.1016/j.orcp.2020.05.005

Wong, G., Greenhalgh, T., Westhorp, G., Buckingham, J., & Pawson, R. (2013). RAMESES publication standards: Meta-narrative reviews. *BMC Medicine, 11*(1), 20. https://doi.org/10.1186/1741-7015-11-20

Wong, G., Greenhalgh, T., Westhorp, G., & Pawson, R. (2014). Development of methodological guidance, publication standards and training materials for realist and meta-narrative reviews: The RAMESES (realist and meta-narrative evidence syntheses - evolving standards) project. *Health Services and Delivery Research, 2*(30), https://doi.org/10.3310/hsdr02300