How does habit form? Guidelines for tracking real-world habit formation

Benjamin Gardner, Amanda L. Rebar & Phillippa Lally |

To cite this article: Benjamin Gardner, Amanda L. Rebar & Phillippa Lally | (2022) How does habit form? Guidelines for tracking real-world habit formation, Cogent Psychology, 9:1, 2041277, DOI: 10.1080/23311908.2022.2041277

To link to this article: https://doi.org/10.1080/23311908.2022.2041277

© 2022 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.
How does habit form? Guidelines for tracking real-world habit formation

Benjamin Gardner¹, Amanda L. Rebar² and Phillippa Lally³*

Abstract: Advances in understanding how habit forms can help people change their behaviour in ways that make them happier and healthier. Making behaviour habitual, such that people automatically act in associated contexts due to learned context-response associations, offers a mechanism for maintaining new, desirable behaviours even when conscious motivation wanes. This has prompted interest in understanding how habit forms in the real world. To reliably inform intervention design, habit formation studies must be conceptually and methodologically sound. This paper proposes methodological criteria for studies tracking real-world habit formation, or potential moderators of the effect of repetition on formation. A narrative review of habit theory was undertaken to extract essential and desirable criteria for modelling how habit forms in naturalistic settings, and factors that influence the relationship between repetition and formation. Next, a methodological

ABOUT THE AUTHOR
Benjamin Gardner is a Reader in Social Psychology at King's College London. His main research interest lies in understanding the concept of habit and how it can be used to explain and change behaviour. He has published >70 papers on habit, variously covering theory development, measurement, behaviour prediction, intervention design and evaluation, and state of the art theoretical and empirical reviews. He co-leads the European Health Psychology Society Habit Special Interest Group. Amanda Rebar is director of the Motivation of Health Behaviours Lab and Associate Professor of Psychology at Central Queensland University in Australia. Her main area of research is on the automatic, irrational motivation of health behaviours and she has a passion for understanding change in behaviour and motivation over time. Phillippa Lally is a Senior Research Fellow in the Research Department of Behavioural Science and Health at University College London. Her research focuses on understanding habitual behaviour and applying this to behaviour change interventions, primarily in the area of energy balance behaviours in those living with and beyond cancer. Her work is currently funded by Yorkshire Cancer Research and Cancer Research UK, and she co-leads the European Health Psychology Society Habit Special Interest Group.

PUBLIC INTEREST STATEMENT
People often act habitually, without thinking beforehand. Psychologists define “habitual” behaviours as actions that are activated automatically when people encounter situations in which they have consistently done that action in the past. It is thought that people tend to sustain habitual actions over time, even if they lose motivation or willpower. Behavioural scientists are increasingly interested in promoting habit formation, to encourage people to adopt and maintain actions that are good for their health or wellbeing. It is important to study how habit forms, to inform such initiatives. This paper offers guidance for behavioural scientists on how to study habit formation in the real-world, and factors that may influence that process. In this paper, we propose criteria for studying habit formation in a way that allows valid conclusions, and show that only five studies to date have met these criteria. We encourage researchers to adhere to these criteria, to increase the likelihood that studies will offer revealing conclusions about how habits develop in real-world settings.
review identified exemplary real-world habit formation studies according to these
criteria. Fourteen methodological criteria, capturing study design (four criteria),
measurement (six criteria), and analysis and interpretation (four criteria), were
derived from the narrative review. Five extant studies were found to meet our
criteria. Adherence to these criteria should increase the likelihood that studies will
offer revealing conclusions about how habits develop in real-world settings.

Subjects: Psychological Science; Testing, Measurement and Assessment; Social Psychology

Keywords: Habit; automaticity; behaviour; behaviour change

1. Introduction

Around half of our everyday behaviours are performed automatically, with little forethought (Wood
et al., 2002). Many such behaviours are instigated by habit, a process whereby situational cues
automatically trigger an impulse to act, via activation of a cue-action association learned through
repetition in the presence of such cues (Gardner, 2015). Actions triggered by this process—i.e.,
“habitual behaviours”—are regulated by contextual stimuli rather than conscious decision-making
processes. Habit plays a vital role in facilitating everyday activity, permitting efficient and effective
initiation of well-rehearsed actions, freeing finite cognitive resources for more challenging con-
current tasks (Neal et al., 2006). This has inspired a growth in research into the role of habit in
everyday settings (Verplanken, 2018). Social psychologists are increasingly applying habit theory to
study real-world behaviours across a diverse range of domains, such as energy conservation
(Walker et al., 2015), education (Hobbs et al., 2021), technology (Limayem & Hirt, 2003), finance
(Allom et al., 2018), health (Gardner, 2015), and intergroup relations (Hackel et al., 2019).

By virtue of its automatic activation, habit increases the likelihood of habitual behaviour being
elicited in associated settings, which translates into more frequent performance over time (Triandis,
1977). Although habitual responses can be inhibited when people have the necessary momentary
motivational and self-regulatory resources needed to act intentionally (Quinn et al., 2010), in the ebb
and flow of daily life, fluctuations in motivation, attention or memory lead habits to tend to dominate
over conscious intentions (Neal et al., 2013; Ouellette & Wood, 1998). Habit can thereby compensate
for momentary dips in the strength of positive intentions. A runner, for example, may go for their
morning run out of habit, despite poor weather having weakened their running intentions (Gardner
et al., 2020). Habit can also lead people to act contrary to intentions. For example, after the 2007
smoking ban in UK public houses, many smokers reported finding themselves lighting a cigarette as
an unintentional habitual response to drinking alcohol (Orbell & Verplanken, 2010).

Habit formation offers a mechanism for behaviour maintenance. Initially successful behaviour
change attempts often fail over the long-term, but theory proposes that a behaviour that becomes
habitual will likely persist even when motivation erodes (Verplanken & Wood, 2006). Commentators
have called for habit formation to be adopted as an intervention goal (Rothman et al., 2009). Understanding how context-consistent behavioural repetition strengthens habit, and factors that reinforce this process, will aid design of habit-forming interventions.

While habit research has traditionally used animal learning paradigms or lab-based tasks in
humans (Smith & Graybiel, 2016), theoretical and methodological advances have inspired growth
in studies of human habit formation for everyday actions in common contexts (Lally et al., 2010;
Shiffman et al., 2008; Verplanken & Orbell, 2003). Research in controlled settings, such as experi-
mental lab-based studies, offers important elucidation of core principles and mechanisms under-
lying habit formation (see, Carden & Wood, 2018). Studying habit formation “in the wild”, however,
can identify impediments that have minimal influence in controlled settings, such as forgetting to
act (Lally et al., 2010), pursuing competing goals (Hamilton et al., 2019), or temporary removal
from target contexts (Lally et al., 2011).
Research questions addressed by real-world habit formation studies fit one of two categories. One focuses on the nature of the relationship between context-consistent behavioural repetition and habit formation; for example, whether each repetition has an equal impact on habit development (Schnauber-Stockmann & Naob, 2018). The other focuses on potential moderators of this relationship, such as whether habit develops more quickly for morning versus evening performances (Fournier, d’Arripe-Longueville, Rovere et al., 2017). These questions can be addressed by two broad study types, discerned by their sensitivity to habit formation parameters. One type involves comparing habit strength across conditions at one or more timepoints, to explore the impact of a given treatment on habit strength. For example, one study focused on whether people instructed to floss daily before versus after brushing differed in flossing habit at 4-week follow-up (Judah et al., 2013). A second study type permits estimation not only of the level at which habit may peak, but the time taken to reach this peak. For example, one study measured repetitions of a dietary or physical activity behaviour and habit strength, and showed that habit peaked at different levels and speeds between participants (Lally et al., 2010). While both study types can generate valid insights into habit formation, the latter, which we term “habit formation tracking studies”, are more informative for intervention development.

Real-world habit formation tracking studies are often resource-intensive, typically requiring multiple assessments, lengthy tracking periods, and complex statistical methods (see, Lally et al., 2010). There is, to our knowledge, no explicit guidance available on how to conduct such studies. Researchers have adopted a variety of methods to investigate human habit formation (see, Gardner & Lally, 2018), and while methodological heterogeneity can enrich understanding, there is a risk that some researchers may unwittingly use methods that give rise to misleading conclusions (A. L. Rebar et al., 2019).

This paper proposes criteria to guide procedures for empirical studies of habit formation in field settings. Our guidance is based on expert opinion: each author has 10–15 years of knowledge and experience accrued from undertaking published social psychology research on habit theory and application. The paper arose from the authors’ experiences, in either invited reviewer or editorial positions at social and health psychology journals, of peer-reviewing a growing body of social psychology studies around real-world habit formation and identifying common limitations that often preclude robust conclusions regarding the habit formation process.

Our criteria, set out in Table 1, offer a practical guide for design and execution of applied social psychology studies investigating the relationship between behavioural repetition and the formation of real-world habits, or moderators of this relationship. Our aim is not to generate or test hypotheses, nor to disparage or discourage studies of habit formation in controlled (e.g., lab-based) settings, but rather to provide methodological support to the growing number of researchers studying the habit formation process “in the wild”. We focus on habit formation tracking studies because these permit exploration of the speed and level at which habit peaks, so are the most informative study designs. Nonetheless, several of our criteria apply to studies that seek to understand the level but not speed at which habit peaks; these criteria are worded to refer to “habit formation studies” in general, rather than “habit formation tracking studies”.

In this paper, we first present a narrative review of theory- and evidence-based principles regarding how habit forms. Using “narrative overview” methods (Green et al., 2001), this review summarises literature in three areas: definitions and operationalisations of habit, the conditions under which habit forms, and the trajectory of habit growth. From this review, we derive criteria for designing studies, measuring habit, and analysing and interpreting data from habit formation studies. While expert opinion renders our guidance subjective, the narrative review serves to justify and explain each criterion. Second, we provide a systematic methodological review of extant habit formation studies. Methodological reviews yield information on the design, conduct or analysis characteristics of extant studies (Mbuagbaw et al., 2020). Using our criteria as a checklist, our review identifies and describes exemplary extant studies that have tracked habit formation or investigated factors that reinforce the impact of repetition on habit development.
2. Identifying criteria for investigating habit formation: A narrative review

2.1. What is “habit”?  
A fundamental criterion for habit formation studies is that they must seek to assess habit to detect to what extent and how it has formed:

<table>
<thead>
<tr>
<th>Domain</th>
<th>Criteria</th>
<th>Essential or desirable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design</td>
<td>D1. Habit formation studies must focus on the strengthening of one or more specific cue-behaviour associations</td>
<td>Essential</td>
</tr>
<tr>
<td></td>
<td>D2. Habit formation studies must be conducted in settings in which there is reason to expect habit to meaningfully strengthen</td>
<td>Essential</td>
</tr>
<tr>
<td></td>
<td>D3. Analyses of potential moderators of habit formation must adopt designs sensitive to the temporal nature of relationships between repetition, habit development, and potential moderators</td>
<td>Essential (for investigating moderators)</td>
</tr>
<tr>
<td></td>
<td>D4. Habit formation studies should use longitudinal designs</td>
<td>Desirable</td>
</tr>
<tr>
<td>Measurement</td>
<td>M1. Habit formation studies must attempt to measure habit</td>
<td>Essential</td>
</tr>
<tr>
<td></td>
<td>M2. Habit formation studies must not infer habit from behavioural frequency alone</td>
<td>Essential</td>
</tr>
<tr>
<td></td>
<td>M3. Habit formation studies must measure habit in relation to a behaviour of interest</td>
<td>Essential</td>
</tr>
<tr>
<td></td>
<td>M4. Habit formation studies should use habit measures relating to behaviour at an appropriate level of specificity</td>
<td>Desirable</td>
</tr>
<tr>
<td></td>
<td>M5. Habit formation studies should use context-specific habit measures</td>
<td>Desirable</td>
</tr>
<tr>
<td></td>
<td>M6. Habit formation studies should measure habit over multiple timepoints</td>
<td>Desirable</td>
</tr>
<tr>
<td>Analysis and Interpretation</td>
<td>AI1. Habit formation studies must acknowledge habit strength as a continuum</td>
<td>Essential</td>
</tr>
<tr>
<td></td>
<td>AI2. Habit formation tracking studies should not infer effects of repetition from measures of time</td>
<td>Desirable</td>
</tr>
<tr>
<td></td>
<td>AI3. Habit formation tracking studies must not assume that habit develops linearly</td>
<td>Essential</td>
</tr>
<tr>
<td></td>
<td>AI4. Habit formation tracking studies must involve analyses that account for individual differences in the growth trajectory</td>
<td>Essential</td>
</tr>
</tbody>
</table>
Measurement Criterion 1 (essential): Habit formation studies must attempt to measure habit.

Yet, there is no consensually agreed definition of habit. Habit has variously been defined or operationalised as a tendency to act (Ouellette & Wood, 1998), a form of automaticity (Verplanken & Wood, 2006), a learned cue-behaviour association (Fleetwood, 2021), or a process by which cued activation of such associations generates a non-conscious impulse to act (Gardner, 2015). All these definitions have attracted criticism (Fleetwood, 2021; Gardner, 2015; Orbell & Verplanken, 2015). Nonetheless, there is consensus around several key features of habit, or habitual responding, with all definitions incorporating a history of repetition, learned cue-behaviour associations, context-consistent performance, and cue-dependent automaticity (Gardner, 2015). It is consensually agreed that habit is neither synonymous with, nor its key features adequately indexed by, past behaviour frequency (Ouellette & Wood, 1998). Inferring habit from repetitive performance is a legacy of behaviourist portrayals of habit as a hypothetical cache of previous performances, with more frequent past performance representing greater potential for subsequent performance (Hull, 1943). It is well-established that behaviour frequency alone fails to distinguish between habit and other processes that generate repeated performance (Ajzen, 2002). Furthermore, theory proposes that habit sustains action (Triandis, 1977), and it is illogical to portray habit as both a form of behaviour and a determinant of behaviour (Maddux, 1997).

Hence:

Measurement Criterion 2 (essential): Habit formation studies must not infer habit from behavioural frequency alone.

Habit tracking studies must specify a behaviour; measures of a generalised tendency to act habitually across behaviours or contexts (e.g., Ersche et al., 2017), for example, will be uninformative when the aim is to track the formation of a specific habit. Hence:

Measurement Criterion 3 (essential): Habit formation studies must measure habit in relation to a behaviour of interest.²

While a dichotomy is often implied between “habitual” and “non-habitual” behaviour, habit is however more appropriately portrayed as a continuum, such that habit becomes stronger with repetition (Moors & De Houwer, 2006). Thus:

Analysis and Interpretation Criterion 1 (essential): Habit formation studies must acknowledge habit strength as a continuum.

Such acknowledgement need not be explicit; simply adopting a continuous measure of habit is sufficient to meet this criterion.

Formation studies must investigate development of features that distinguish more strongly habitual actions from weaker or non-habitual actions. Habit formation is essentially a process of learning new cue-behaviour associations, which gradually acquire the potential to activate impulses to act when triggered by exposure to the cue. By definition, then:

Design Criterion 1 (essential): Habit formation studies must focus on the strengthening of one or more specific cue-behaviour associations.

While cue-behaviour associations cannot be directly observed, multiple measures are available to estimate associations from their sequelae. The most direct assessments infer the strength of cue-behaviour associations from observed responses to hypothesised cues. These association-based measures operate on the basis that, in the presence of cues, habitual responses are most cognitively accessible so will be exhibited more quickly, frequently or accurately (Danner et al., 2011), and have been suggested to be the “gold standard” of habit measurement (Gardner, 2015).
Such measures infer strong habit from quicker reaction times or higher frequency with which hypothesised cues elicit expected responses (Verplanken et al., 1994), or from impaired reaction times to tasks incompatible with hypothesised responses (Luque et al., 2020).

Self-report measures indirectly estimate the strength of cue-behaviour associations. “Frequency in Context” (FiC) measures focus on the context-specificity of behavioural repetition, so assess the likelihood that habit will have developed based on the consistency of performance contexts (Labrecque & Wood, 2015). However, FiC measures are incompatible with designs that track the contribution of behavioural repetition to habit formation: behaviour frequency should not conceptually or statistically be modelled as both a precursor of habit formation and as part of a habit strength measure.

Non-frequency-based self-report measures offer distal assessment of cue-behaviour associations, by eliciting reflections on “symptoms” of habitual responding, such as acting without awareness (Orbell & Verplanken, 2015). Some such measures have been designed for particular behavioural domains, such as physical activity (Tappe & Glanz, 2013) or use of information technology (Limayem & Hirt, 2003). Others are designed for adaptation to any behaviour. The generic Self-Report Habit Index (SRHI; Verplanken & Orbell, 2003), for example, comprises statements about experiences of habitual action, with which participants rate agreement (Orbell & Verplanken, 2015). SRHI items follow a stem (“[Behaviour X] is something …”) and tap perceptions of automaticity (“… I do without thinking”), repetition frequency (“… I do frequently”) and history (“… I have been doing for a long time”), and identity congruence (“… that is typically ‘me’”; Verplanken & Orbell, 2003). The SRHI has however been criticised on the basis that, while self-identity is often associated with habit, it is not a necessary component of habit (Gardner et al., 2012), and cannot be expected to develop in concert with habit. Additionally, as with FiC measures, the inclusion of behavioural repetition items renders the SRHI problematic when modelled as an outcome of repeated performance. The Self-Report Behavioural Automaticity Index (SRBAI; Gardner et al., 2012), a derivative of the SRHI, avoids this limitation by focusing only on automaticity (but see, Keatley et al., 2015). Conversely however, debate surrounds the conceptual validity of the SRBAI because it omits items assessing repetition history, so fails to distinguish between habit-related automaticity (acquired through context-dependent performance) and non-habit-related automaticity (e.g., acquired through mental rehearsal; Orbell & Verplanken, 2015). However, studies of habit formation over time inherently account for repetition, rendering this reservation irrelevant in habit tracking studies.

Regardless of which measure is adopted, it would seem prudent to measure habit development for specific behaviours, performed in specific contexts. SRHI and SRBAI items, for example, can be adapted to incorporate contextual cues (“Behaviour X in Context Y is something I do without thinking”; Sniehotta & Presseau, 2011). Many studies have assessed habit in relation to broad behavioural categories (e.g., “eating two pieces of fruit per day”; Brug et al., 2006), or used context-free measures (e.g., “smoking is something I do automatically”; Kovač & Rise, 2008). Such measures fail to distinguish between closely related behaviours that may differ in habit strength (e.g., habitually eating a banana with breakfast, but non-habitually eating an apple after dinner), or contexts in which the behaviour is performed habitually and those in which it is not (e.g., smoking habitually upon waking, but consciously deciding to smoke when invited by a friend). In theory, such measures should summarise habit across behaviours or contexts, such that, all else being equal, gains for a specific behaviour in a specific context should translate into increased aggregate habit strength. While suboptimal, broad behaviour measures should therefore be adequate to detect changes in habit over time, so we do not deem specific behaviour measures to be essential. Greater specificity in behaviour descriptions is not however uniformly more preferable; people are unlikely to conceive of the act of eating an apple according to the fine-grained muscle movements required to chew, for example, (Vallacher & Wegner, 1987). Optimal behaviour measures will likely correspond most closely with the level at which people typically construe their own actions. Thus:

Measurement Criterion 4 (desirable): Habit formation studies should use habit measures relating to behaviour at an appropriate level of specificity.
Measurement Criterion 5 (desirable): Habit formation studies should use context-specific habit measures.

The latter criterion is not deemed essential because, for behaviours that can realistically only be performed in one context (using a continuous positive airway pressure machine while in bed; Broström et al., 2014), including context cues in habit measures may be unnecessary.

In sum, habit is a determinant of behaviour, not a type of behaviour. While a consensual definition of habit per se remains elusive, there is agreement that habitual behaviour arises from cued activation of cue-behaviour associations acquired through repeated performance. Habit formation studies must attempt to capture the strengthening of associations between context cues and behavioural responses.

2.2. In what conditions does habit form?
Habit formation studies must only be conducted when habit is expected to form. Habit develops when a behaviour is repeated in a consistent context (Fournier, d’Arripe-Longueville, Rovere et al., 2017; Lally et al., 2010). Increases in behavioural repetition must be brought about by changes to one or more of three fundamental determinants: motivation, capability, and opportunity (Michie et al., 2011). Changes in motivation, capability or opportunity may be purposefully brought about by the individual, by: forming an intention to take up a new action (e.g., resolving to exercise every morning), which enhances motivation; mobilising self-regulatory resources (e.g., watching instructional exercise videos), which increases the capability to act; or seeking or creating new possibilities for action (e.g., waking up earlier than normal, to allow more time to exercise), which boosts opportunity. Interventions may be developed to modify motivation or capability, or enable or restrict opportunities, to promote context-dependent repetition. For example, in a study of oral hygiene habit development, participants were asked to form implementation intentions to identify cues to once-daily flossing, to enhance capability and so increase the likelihood of action (Judd et al., 2013). Alternatively, purposeful or naturally occurring changes to everyday contexts may spontaneously yield habit formation opportunities. Such changes may be temporary; for example, road closures restrict opportunities to drive, which can encourage car commuters to adopt new public transport habits (Brown et al., 2003). Others may be longer-lasting; major life events, such as becoming a parent, moving home, or a workplace relocation create opportunities for people to develop new cue-behaviour associations (Verplanken et al., 2018).

Regardless of why such changes may occur, habit formation must be preceded by changes to perceived or actual motivation, capability or opportunity to act, in order to trigger early behavioural performances (Lally & Gardner, 2013); in the absence of such changes, habit will not form. Thus:

Design Criterion 2 (essential): Habit formation studies must be conducted in settings in which there is reason to expect habit to meaningfully strengthen.

Tracking habit scores over time will not capture meaningful habit formation without a naturally occurring or purposeful change in the determinants of context-consistent repetition. For example, predictive models of habit become de facto analyses of habit change by statistically controlling for habit strength at an earlier timepoint. Yet, if there is no reason to expect habit to have changed over the intervening period, variation in the outcome cannot reliably be attributed to meaningful habit change. Such variation may instead reflect natural fluctuations, or methodological error arising from inconsistent responding over multiple timepoints (A. L. Rebar et al., 2019). Notably few habit measures have been assessed for test-retest reliability or sensitivity to change (but see, Verplanken & Orbell, 2003).

By extension, attempts to investigate determinants of habit formation will yield minimal insight where habit has not reliably changed. Prospective studies of established, ongoing behaviours, for which habit strength will likely have plateaued (Lally et al., 2010), cannot reveal the causal relationships between variables that preceded development of habit (Weinstein, 2007). For
example, one study found that fruit consumption frequency was a stronger predictor of fruit consumption habit, as measured four weeks later, among people who at baseline reported finding fruit consumption more intrinsically rewarding than those who did not (Wiedemann et al., 2014). It would be inappropriate to classify these findings as evidence of the reinforcing effect of intrinsic rewards on habit development, as they may mask a complex web of bidirectional relationships between variables that unfolded prior to the study being conducted. For example, as habit formed and behavioural initiation became easier, the behaviour may have become more intrinsically rewarding, prompting more frequent performance, and further strengthening habit.

True accounts of the impact of antecedents on habit development require longitudinal designs that track changes in habit strength in accordance with the hypothesised temporal sequence of such relationships. Hence:

Design Criterion 3 (essential; moderator studies only): Studies of potential moderators of habit formation must adopt designs sensitive to the temporal nature of relationships between repetition, habit development, and potential moderators.

Studies of established habits and their correlates contribute to understanding by generating hypotheses for investigation in settings in which habit is expected to form.

We do not offer guidance on which types of behaviour should be the focus of habit formation studies, because, in theory, any behaviour, when sufficiently repeated in a given context, can become habitual (Verplanken, 2005; but see, Mullan & Novoradovskaya, 2018). Habit formation need not involve uptake of new behaviours; habit can develop when familiar behaviours are performed in new contexts. Yet, targeting habit formation for familiar behaviours can be challenging. People who volunteer for behaviour change studies often have a strong prior interest in the target behaviour, and likely already perform it (De Bruin et al., 2015), which may limit further habit development. For example, in a trial of an intervention designed to promote formation of physical activity habits among highly inactive older adults, participants reported over 30 mins of average daily activity at baseline, far exceeding the 150 mins of weekly activity recommended by international guidelines (White et al., 2017). Perhaps consequently, habit strength—as assessed via agreement with a single SRHI item (“physical activity is something I do automatically”)—did not appear to change meaningfully in intervention recipients, nor in an active control group. While it is possible that the broad behavioural label (“physical activity”) or context-free measure obscured detection of recommended habit-forming activities (e.g., heel raises while washing dishes), participants may alternatively have already had settled patterns of context-dependent physical activity engagement, limiting the capacity for further habit development. These possibilities speak to the importance of studying formation among people with sufficient capacity for habit to develop for a given behaviour in a given context.

While it is possible for people to form habits by performing familiar behaviours more consistently, the “purest” account of the formation process might be obtained by focusing on wholly unfamiliar actions, for which a complete absence of habit can be reliably assumed. One innovative example sought to instil the habit of microwaving sponges, a little-known food hygiene behaviour (Mullan et al., 2014). At baseline, none of the 45 study participants engaged in the behaviour, which was reflected in minimal habit scores. Observed increases in habit strength could therefore be reliably attributed to exposure to the intervention.

In sum, habit meaningfully develops in response to enhancements in the motivation, capability or opportunity required to repeatedly act in a stable context. Habit formation studies must be undertaken only where there is reason to believe such enhancements have occurred.

2.3. How does habit form?
Adequately tracking the strengthening of habit requires methods sensitive to the habit formation trajectory. This requires decisions around the number of measurement points over which habit
should be tracked. Habit formation can be inferred from single-timepoint data obtained after motivational and contextual changes have triggered habit development, but only reliably so for behaviours in which participants have not previously engaged (Mullan et al., 2014). However, where target behaviours have been performed previously, repeated assessments are needed to reveal changes in habit strength from a baseline value. Thus, we recommend that:

2.4. Design Criterion 4 (desirable): Habit formation studies should use longitudinal designs

While habit necessarily forms over time, the key determinant of habit development is consistent repetition, not time; habit would be expected to strengthen over a shorter time period for an action performed on multiple occasions daily than for a once-weekly activity. For behaviours performed once-daily, the distinction between repetition and time becomes less relevant, because time can be treated as a proxy for repetition. However, the two variables are not wholly interchangeable; even for once-daily activities, plotting habit over time will overlook missed performances. Hence:

Analysis and Interpretation Criterion 2. (desirable): Habit formation tracking studies should not infer effects of repetition from measures of time.

Several studies have found habit to strengthen asymptotically over time, with rapid early gains slowing and levelling off (Fournier, d’Arripe-Longueville, Rovere et al., 2017; Lally et al., 2010; but see, Keller et al., 2021; Schnauber-Stockmann & Naab, 2018). Asymptotic growth renders the use of linear analyses to model habit formation problematic. For example, using linear regression analysis to model the effect of repetition frequency on habit development may erroneously assume that each repetition makes an equal contribution to habit formation, and that habit can strengthen indefinitely. Similarly, a two-timepoint design can detect habit development, but cannot distinguish between linear and non-linear growth.

There has been disagreement around how best to depict the habit growth trajectory. Lally et al. (2010) portrayed habit formation using an exponential asymptotic growth curve, with habit strengthening immediately in response to initial repetitions, and slowing as a plateau was reached. Fournier, d’Arripe-Longueville, Rovere et al. (2017), however, depicted formation using a logistic function, with earliest repetitions expected to have negligible impact, followed by a period of pronounced, rapid growth, which decelerated as habit plateaued. However habit development is modelled, estimating a non-linear growth curve requires at least three measurement points. Still more intensive measurement schedules are however recommended, as they provide a richer insight into the progress of habit development. For example, they can differentiate between momentary fluctuations in habit strength versus more meaningful longer-term development (Conroy et al., 2013). More measurement points allow more precise estimation of core parameters such as the level at which habit peaks for a given individual performing a given behaviour in a given context, the rate at which habit develops, the time taken for habit to peak, and more broadly, the fit of a hypothetical growth curve to the observed data (Fournier, d’Arripe-Longueville, Rovere et al., 2017; Lally et al., 2010). Insight into potential moderators of the impact of repetition on habit formation can be gleaned from comparison of the impact of such variables on these parameters (Schnauber-Stockmann & Naab, 2018). For example, the potential moderating effect of intrinsic motivation on the relationship between context-dependent repetition and habit formation can be investigated by comparing the level at which habit peaks, and the time taken to reach peak habit strength, between actors identified a priori as intrinsically motivated versus those not intrinsically motivated to act (see, Gardner & Lally, 2018).

Habit may develop linearly among some individuals, behaviours or contexts (Schnauber-Stockmann & Naab, 2018). However, evidence of asymptotic growth falsifies the assumption that habit growth is necessarily linear. Thus:
Analysis and Interpretation Criterion 3 (essential): Habit formation tracking studies must not assume that habit develops linearly.

Habit development is inherently idiosyncratic, based on personally acquired behavioural responses to personally relevant cues. The habit formation trajectory is determined by interplay between characteristics of the target behaviour, performance context, and actor. Formation studies must therefore capture habit development within individuals. Aggregating across individuals will obscure individual-level variation in habit formation. Hence:

Analysis and Interpretation Criterion 4 (essential): Habit formation tracking studies should involve analyses that account for individual differences in the growth trajectory.

Individual-level analysis methods are available: N-of-1 designs, for example, capture variation within individuals over time. Modelling effects at multiple levels may be most informative; for example, Lally et al. (2010) undertook within-person analyses of the effect of missed performances on the habit trajectory, and compared patterns between participants to identify generalities, such as the typical shape of the habit growth curve.

In sum, studies must recognise that habit forms with multiple cue-behavioural response repetitions, and may follow a non-linear trajectory.

3. Exemplary habit formation tracking studies: A methodological review

Table 1 summarises the nine essential and five desirable criteria conducive to robust conclusions in studies tracking the impact of repetition on habit formation, and moderators of such impact. To illustrate exemplary practice according to our criteria, we undertook a methodological review to identify and describe methodological characteristics of extant quantitative studies of real-world human habit formation that met all applicable essential criteria (i.e. exemplary habit formation tracking studies; Mbuagbaw et al., 2020). Although synthesis of study findings was outside the scope of the review, for the sake of completeness a narrative summary is presented as Supplementary Text.

A systematic literature search of four psychology and health databases (Embase, Medline, PsycInfo, Web of Science) was undertaken in April 2021. A cited reference search was used to identify studies that had cited one or more key papers in habit measurement (i.e., Gardner, Abraham et al., 2012; Ouellette & Wood, 1998; Verplanken & Orbell, 2003), principles and processes of habit formation (i.e., Gardner, Lally et al., 2012; Lally & Gardner, 2013; Lally et al., 2010, 2011), and conceptual commentaries (i.e., Gardner, 2015; Wood & Rünger, 2016). Studies were eligible where they reported primary analyses of quantitative data from human participants regarding the process of forming real-world habits, were published in English in a peer-reviewed journal, and met criteria D1-2, M1-3, and A11, A13 and A14. Investigations of moderators were also required to meet criterion D3.

A “real-world” habit was defined as pertaining to a behaviour that a person could realistically be expected and motivated to learn and repeatedly perform outside of a controlled research setting (e.g., physical activity; dietary consumption), in contrast with more artificial actions or contingencies created or performed solely in controlled research settings with minimal external application (e.g., button-pressing; Smeets et al., 2019). Studies of the feasibility, acceptability or effectiveness of “habit formation” as a behaviour change technique (e.g., Hamilton et al., 2019) were excluded unless they also provided data relating to the process via which habit formed, in accordance with our essential criteria. While there was no explicit requirement for a minimum number of time-points within any reviewed study (see desirable criterion M6), adherence to the non-linearity criterion (A13; essential) required that studies assessed habit on three or more occasions.

Of 6466 records, 3150 were duplicates. Screening titles and abstracts removed 3028 obviously irrelevant papers, and full-text inspection excluded a further 283 records. Several studies were excluded solely because they used linear models, so violating the non-linearity criterion (A13; e.g., Fournier,
Five papers each reported a single study, which met all eligibility criteria (see Supplementary Table). These studies focused on dietary behaviour (N = 117; Keller et al., 2021), physical activity (a stretching exercise; N = 42; Fournier, d’Arripe-Longueville, Rovere et al., 2017), dietary behaviour and physical activity (a participant-chosen behaviour; N = 96; Lally et al., 2010), media use (use of a football tournament smartphone app; N = 51; Schnauber-Stockmann & Naab, 2018), and, in one study, a self-chosen behaviour from a range of options relating to health, interpersonal relationships, finance, or pro-environmental behaviours (N = 146; Van der Weiden et al., 2020). All five both tracked habit formation and sought to explore potential moderators of the impact of repetition on formation, taken over periods of 30 days (Schnauber-Stockmann & Naab, 2018), 84 days (Keller et al., 2021; Lally et al., 2010), and 90 days (Fournier, d’Arripe-Longueville, Rovere et al., 2017; Van der Weiden et al., 2020; D3, D4). Two studies sought to compare linear and non-linear accounts of the relationship between repetition and habit development, and two studies assumed a priori that habit forms non-linearly (Fournier, d’Arripe-Longueville, Rovere et al., 2017; Keller et al., 2021; A13). One study used non-linear, asymptotic growth curve modelling at the individual level in accordance with criterion A13, with linear modelling at the group level to model moderators of habit formation (Van der Weiden et al., 2020), in violation of A13.

Three studies explored the impact of missed performances on habit development. Moderators investigated were behaviour type, cue type, time of day (morning or evening), context stability, self-control capacity and perceived reward value. All studies focused on researcher-prompted habit formation attempts, with participants forming action plans to initiate action in response to a target context encountered once daily, either selected by a researcher (e.g., “upon waking” or ‘before sleeping’; Fournier, d’Arripe-Longueville, Rovere et al., 2017); event-based vs time-based cue (Keller et al., 2021) or the participant (Lally et al., 2010; Schnauber-Stockmann & Naab, 2018; Van der Weiden et al., 2020; D1, D2). All studies assessed behaviour frequency—in four studies, whether the behaviour was performed on the target day, and in one study, whether and how often the behaviour was performed—and habit, using the SRHI or SRBAI, via daily self-report (A11, M1, M2, M3, M6). In one study, the habit measure specified each participant’s chosen behaviour and chosen context (Lally et al., 2010), and in another, while the measure assessed only behaviour (“performing my stretching exercise”), participants were expected to act only at the specified time of day (Fournier, d’Arripe-Longueville, Rovere et al., 2017; M3, M4, M5). Three studies used measures specifying a behaviour (e.g., “today I opened the app”) but no context (Keller et al., 2021; Schnauber-Stockmann & Naab, 2018; Van der Weiden et al., 2020; M3, M4). All studies used both within-participant analyses to track habit formation (A14), and comparisons across participants to assess moderators. All studies analysed habit formation over time, rather than repetition (contrary to criterion A12). In one study, participants were found to perform the behaviour more than once daily, problematising use of time as a repetition index (Schnauber-Stockmann & Naab, 2018).

4. Discussion
Recent years have seen growing interest among social psychologists in understanding how habit forms and what might aid formation, with a view to applying such insights to promote sustainable positive behaviour change in a range of behavioural domains (Verplanken, 2018). Drawing on theory- and evidence-based principles regarding what habit is, and how it develops, we proposed 14 design, measurement, analysis and interpretation criteria for quantitative studies tracking the habit formation process and contributory factors in real-world settings. Applying criteria deemed essential for robust habit formation tracking studies, we identified five exemplary extant studies. Future discoveries in the science of human habit formation may prompt refinement of our criteria. We encourage researchers to adhere to these criteria when planning real-world habit formation studies, to increase the likelihood that such discoveries are made.
While this paper was not designed to generate new research questions, our methodological review of exemplary applied studies to date nonetheless revealed important gaps in applied habit formation research. Only five studies met our criteria. All focused on behaviours performed at least once daily. More work is needed to track formation for infrequently but consistently performed actions. Additionally, the five studies focused on researcher-initiated, rather than spontaneous participant-led, habit formation attempts. Future research might explore how people form new habits in response to major context changes that discontinue old habits (i.e. habit discontinuity; Verplanken et al., 2018), or when new responses compete directly with existing cue-responses (i.e. habit substitution; Gardner, Rebar et al., 2020). We believe that following our criteria will enhance the validity of studies addressing important applied habit research questions. For example, in light of disagreements of the conduciveness of complex behaviours to habit formation, a recent agenda for applied habit research within health psychology highlighted the need to understand the extent to which behavioural complexity influences habit formation (Gardner et al., 2021). Adhering to our criteria would facilitate the design and execution of studies tracking potentially non-linear, within-participant development of habits pertaining to behaviours of various complexity, and comparison of formation trajectories according to complexity, over time.

While adhering to our criteria may enhance the robustness of study conclusions, this is not to say that studies that have not met our criteria have yielded invalid findings. For example, any observation of habit change in real-world settings in which habit can be expected to form constitutes a “real-world habit formation study”, regardless of whether our essential analysis and interpretation criteria are met. While studies that investigate both the extent of habit gains and the speed with which such gains are made are optimal for understanding habit development, studies that violate our criteria by focusing only on the magnitude of changes in habit can nonetheless be informative where they adhere to our essential design and measurement criteria. For example, Judah et al. (2013) showed newly-adopted dental flossing habit scores to decline between one month and six months, so demonstrating that, contrary to theoretical predictions (e.g., Verplanken & Wood, 2006), habit gains may be lost over time. Similarly, an investigation of development of child-feeding habits (providing fruit, vegetables, or water), demonstrated that parents of 2–6 year olds were able to form equally strong habits for the three behaviours in quick succession (Gardner et al., 2014). Yet, other conclusions reached by such studies—for example, that flossing after brushing is more conducive to habit formation (Judah et al., 2013), or that action plan specificity has no bearing on habit formation (Gardner et al., 2014)—may be less well-founded because, in violation of our non-linearity criterion, they arose from linear modelling analyses, which may inadequately depict true formation trajectories. While our proposed design and measurement criteria apply to all real-world habit formation studies, we recommend adherence to all our essential criteria, to ensure all study findings are robust.

Our criteria may seem overly restrictive. For example, by cautioning against using “frequency in context” measures to capture the relationship between repetition and habit formation, we discount the use of “big data” technologies to model habit development based on emergence of observable, context-consistent behavioural patterns (see, Carden & Wood, 2018). However, such studies track only predictable behaviour, not habit per se, and it is problematic to treat behaviour as both a precursor of habit and an index of the habit so generated. Contextually stable behavioural performance may arise from processes other than habit (Volpp & Loewenstein, 2020). “Big data” cannot be treated as definitive observations of habit formation.

We recognise limitations of our proposed criteria. The criteria were derived from and so are constrained by current theory and evidence. Many of them therefore represent suggestions for avoiding bad practice rather than achieving best practice. For example, the lack of a consensually agreed definition of “habit” renders criteria for assessing habit development inherently problematic. While there is general consensus around key features of habit or habitual responding (Gardner, 2015), debate remains around the optimal measure of real-world habits (Hagger et al., 2015; Labrecque & Wood, 2015; Orbell & Verplanken, 2015). While we urge researchers to eschew solely frequency-
based habit measures, we cannot identify the “best” habit measure to use (for a review, see, A. Rebar et al., 2018). At a minimum, we recommend that researchers adhere to our measurement criteria by making explicit their definition of habit, selecting a habit measure that most closely fits their definition, and acknowledging the fit between definition and measure. With regards to analysis, we discourage use of linear analysis methods, but ongoing debate surrounding non-linear alternatives prevents us from recommending how best to model non-linear habit growth (e.g., Fournier, d’Arripe-Longueville, Rovere et al., 2017; see, Gardner & Lally, 2018, for practical guidance for one non-linear modelling option). More work is needed to identify conceptually and statistically optimal within-participant habit modelling techniques. More broadly, habit theory and methods must develop further before guidelines for best practice for modelling habit formation can credibly be offered.

Our criteria are based on expert opinion, so are subjective. For example, we recommend that behaviour is measured at an appropriate level of specificity. While our aim is to discourage use of broad behavioural categories (e.g., “physical activity”) to measure habitual performance of specific behaviours (e.g., “doing a stretching exercise”; Fournier, d’Arripe-Longueville, Rovere et al., 2017), it is difficult to identify the most appropriate level at which to portray behaviour. Any one behaviour can be mentally represented at multiple levels of analysis—for example, “doing a stretching exercise” can be represented at a finer-grained level (e.g., “moving my arm up”), or more abstract levels (e.g., “being physically active”). Representations vary between people, and within a person over time (Vallacher & Wegner, 1987). Some argue that behaviour is most appropriately represented to accord with participants’ own representations (Gardner et al., 2019), but this is an empirical question.

Promoting habit in real-world settings requires understanding how habit forms, and how to quicken habit formation, or achieve stronger habits. We have proposed criteria for studies of real-world habit formation. Adhering to these criteria will enhance the likelihood that future applied psychology studies will generate robust conclusions around how habit forms.

Funding
The authors received no direct funding for this research.

Author details
Benjamin Gardner1
E-mail: benjamin.gardner@kcl.ac.uk
ORCID ID: http://orcid.org/0000-0003-1223-5934
Amanda L. Rebar2
E-mail: a.rebar@cqu.edu.au
ORCID ID: http://orcid.org/0000-0003-3164-993X
Phillippa Lally3
E-mail: p.lally@ucl.ac.uk
1 Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, UK.
2 Physical Activity Research Group, Appleton Institute, School of Health, Medical, and Applied Sciences, Central Queensland University, Rockhampton, Australia.
3 Research Department of Behavioural Science and Health, University College London, London, UK.

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

Data availability statement
Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Supplementary material
Supplemental data for this article can be accessed here.

Citation information
Cite this article as: How does habit form? Guidelines for tracking real-world habit formation, Benjamin Gardner, Amanda L. Rebar & Phillippa Lally, Cogent Psychology (2022), 9: 2041277.

Notes
1. A distinction has been drawn between habitually instigated behaviour (or “habitual deciding”), whereby habit bypasses the decision-making process and automatically generates a commitment to act in the absence of forethought, and habitually executed behaviour (or “habitual doing”), whereby habit facilitates progression through the sequence of lower-level acts required to complete any one action (Gardner, Phillips & Judah, 2016). Most social psychology research into “habitual behaviour” focuses on habitual instigation, not execution (Gardner et al., 2019). We use the terms “habit” and “habitual behaviour” hereon to denote habitual instigation and habitually instigated behaviour only. Habitual execution is not discussed further.
2. We recognise that this criterion incorporates, and so renders redundant, Measurement Criterion 1. Nonetheless, we present Measurement Criteria 1 and 3 as separate to emphasise the two distinct points that justify the criteria, namely the importance of measuring habit (Measurement Criterion 1), and the importance of relating habit to a given behaviour (Measurement Criterion 3).
3. Habit forms on the basis of repeatedly acting following exposure to a cue. While the term “before sleeping” suggests that the target act was proposed to precede a cue (“sleeping”), we interpreted “before sleeping” to represent a personally-relevant event that reliably precedes the act of sleeping, so could feasibly cue the target behaviour. Many people engage in a consistent sequence of acts before sleeping (e.g., Quante et al., 2019), making it feasible for a new behaviour to be inserted into an existing sleep hygiene routine (Judah et al., 2013).
4. The definition adopted when undertaking this work was of habit as a process, whereby exposure to a cue activates learned cue-behaviour associations, which in turn
activate impulses to enact the learned response (Gardner, 2015). Our recommendations are however compatible with portrayals of habit as an association (Fleetwood, 2021), a form of automaticity (Verplanken & Wood, 2006), or a tendency (Ouellette & Wood, 1995), all of which distinguish habit as a determinant from the behaviour that it so generates (Maddux, 1997).

References


