

European Journal of Social Psychology
Eur. J. Soc. Psychol. **40**, 998–1009 (2010)
Published online 16 July 2009 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/ejsp.674

Research article

How are habits formed: Modelling habit formation in the real world[†]

PHILLIPPA LALLY*, CORNELIA H. M. VAN JAARVELD,
HENRY W. W. POTTS AND JANE WARDLE
University College London, London, UK

Abstract

To investigate the process of habit formation in everyday life, 96 volunteers chose an eating, drinking or activity behaviour to carry out daily in the same context (for example 'after breakfast') for 12 weeks. They completed the self-report habit index (SRHI) each day and recorded whether they carried out the behaviour. The majority (82) of participants provided sufficient data for analysis, and increases in automaticity (calculated with a sub-set of SRHI items) were examined over the study period. Nonlinear regressions fitted an asymptotic curve to each individual's automaticity scores over the 84 days. The model fitted for 62 individuals, of whom 39 showed a good fit. Performing the behaviour more consistently was associated with better model fit. The time it took participants to reach 95% of their asymptote of automaticity ranged from 18 to 254 days; indicating considerable variation in how long it takes people to reach their limit of automaticity and highlighting that it can take a very long time. Missing one opportunity to perform the behaviour did not materially affect the habit formation process. With repetition of a behaviour in a consistent context, automaticity increases following an asymptotic curve which can be modelled at the individual level. Copyright © 2009 John Wiley & Sons, Ltd.

Performing an action for the first time requires planning, even if plans are formed only moments before the action is performed, and attention. As behaviours are repeated in consistent settings they then begin to proceed more efficiently and with less thought as control of the behaviour transfers to cues in the environment that activate an automatic response: a habit. How long does it take to form a habit? This question is often asked by individuals who want to acquire healthy habits or those who want to promote behaviour change. However, we are not aware of any studies that have systematically investigated the habit formation process within individuals, and none have examined the development of 'real world' habitual behaviours.

The past decade has seen a resurgence of interest in the topic of habitual behaviour within social psychology. There is still debate over how habits should be conceptualized and operationalized, but there is consensus that habits are acquired through incremental strengthening of the association between a situation (cue) and an action, i.e. repetition of a behaviour in a consistent context progressively increases the automaticity with which the behaviour is performed when the situation is encountered (Verplanken, 2006; Wood & Neal, 2007). 'Automaticity' is evidenced by the behaviour displaying some or all of the following features: efficiency, lack of awareness, unintentionality and uncontrollability (Bargh, 1994).

Historically, habit has been measured by assessing the frequency of past behaviour (Sheeran, 2002; Ajzen, 2002). However, this assumes that the relationship between repetition and automaticity is linear, which is unlikely because it implies that every repetition, whether early on in the habit formation process or after a large number of repetitions, would result in the same increase in automaticity. More plausibly, a behaviour becomes as automatic as it can get after some

*Correspondence to: Philippa Lally, Cancer Research UK Health Behaviour Research Centre, Department of Epidemiology and Public Health, University College London, Gower Street, London, WC1E 6BT, UK. E-mail: p.lally@ucl.ac.uk

[†]This research was conducted by Philippa Lally when she held a Medical Research Council PhD studentship and has been written up during an Economic and Social Research Council postdoctoral fellowship.

number of repetitions, and further repetition no longer increases automaticity. In addition, early repetitions could be expected to result in larger increases as the association between situation and action is created.

Hull's early work (1943, 1951) suggested that the relationship between repetition and habit strength follows an asymptotic curve in which automaticity increases steadily—but by a smaller amount with each repetition—until it reaches an asymptote (plateau). This idea was based on laboratory data from human and animal studies where habit strength was inferred from behavioural responses (e.g. drops of saliva secreted in response to a cue) rather than the subjective experience of automaticity. However, because outcomes were averaged across groups, it was not clear whether the curve applied at an individual level.

The exact nature of the relationship between repetition and automaticity, and the time it takes for habit acquisition to occur, have been relatively unexplored because the recent literature has focused largely on established habits rather than the *process* of habit development. There is a clear need to extend Hull's work by examining individual responses and real-world behaviours. This idiographic approach allows conclusions about the general process of habit formation to be drawn from examination of the process across individuals.

THE PRESENT RESEARCH

The aim of the present study was to investigate the development of automaticity in a series of volunteers who were asked to repeat a behaviour of their choice, in response to a cue, in an everyday setting, and complete a measure of automaticity on a daily basis. As described above, it is generally agreed that repeating a behaviour in a consistent situation allows cue-response links to be formed. We therefore tried to control context stability by asking participants to carry out the behaviour in the same situation each day. We selected situations rather than times as cues, because prospective memory research indicates that situations permit external cueing of an intended action whereas time cues require monitoring to identify the time to act (McDaniel & Einstein, 2000). We used daily automaticity scores as the outcome to model increases in automaticity over time within individuals, and assess the time taken for each participant to reach their maximum automaticity.

The role of reinforcement/rewards in habit development and maintenance is unclear. In the behaviourist tradition, habit strength has been considered to be a function of repetition only when rewards are received for performing the behaviour upon encountering a cue (Hull, 1943, 1951); without reinforcement there can be no habits. However, it is unclear whether external rewards need to be provided after each repetition if the behaviour is inherently rewarding. In the present study, participants chose behaviours they wanted to perform; we assumed these to be intrinsically rewarding and no extrinsic rewards were given for performing the behaviour.

In the habit literature, three different methods have been used to quantify habit strength in humans, although they have predominantly been used to characterize existing behaviours as habitual or not, rather than to examine the *process* of habit acquisition. The response frequency measure (RFM) (Verplanken, Aarts, van Knippenberg, & van Knippenberg, 1994) assesses generalized habits across different situations (e.g. car use when travelling to different locations) and so is not applicable to assessing change in a specific behaviour in a specific situation. A second approach (Wood, Quinn, & Kashy, 2002; Wood, Tam, & Witt, 2005) is derived from multiplying ratings of the frequency of performance of the behaviour by the consistency of performance context. Because this is based on an assessment of behavioural frequency, it does not lend itself to tracking habit formation in response to deliberate daily repetition. The third measure is the self-report habit index (SRHI) (Verplanken & Orbell, 2003), which assesses a respondent's subjective experience of several features of habits, three of which are seen as central to automaticity (lack of awareness, lack of control and efficiency), along with two others (history of repetition and identity), in a 12-item questionnaire. The SRHI has been shown to have good test-retest reliability, to correlate well with the RFM and past behaviour measures, and can differentiate between behaviours that are performed with different frequencies (Verplanken & Orbell, 2003). The SRHI has discriminant validity over and above past frequency of behaviour in predicting future behaviour. It has also been used to show that for the same number of repetitions a simple behaviour has a higher habit score than a complex behaviour (Verplanken, 2006).

Habit strength measured using the SRHI (or a subset of SRHI items) has been shown to be associated with various eating behaviours including snacking and fruit and vegetable consumption in adults, and sugar-sweetened beverage consumption in adolescents (Brug, de Vet, de Nooijer, & Verplanken, 2006; Verplanken & Orbell, 2003; Kremers, van der Horst, & Brug, 2007; Verplanken, 2006; Reinaerts, de Nooijer, Candel, & de Vries, 2007; de Bruijn, Kremers, de Vet, de

Nooijer, van Mechelen, & Brug, 2007). Habit strength has also been shown to be predictive of physical activity in adults (Chatzisarantis & Hagger, 2007; Verplanken B & Melkevik O, 2008) and children (Kremers & Brug, 2008). Although it has been suggested that the SRHI could be used to assess change in habit strength over time (Verplanken, Myrbakk, & Rudi, 2005), we are not aware of studies that have used it in this way. For the present study, we generated an automaticity sub-scale which excluded the frequency/history of repetition items because frequency was pre-defined in the study instructions and repetition was built into the intervention. We also excluded the identity items because of controversy over whether habits are more or less fundamental to an individual's identity than non-habits (Verplanken & Orbell, 2003; Wood et al., 2002).

Verplanken's (2006) findings have suggested that complex behaviours achieve lower levels of automaticity than simple behaviours. However, in his study, the number of repetitions was restricted, leaving it possible that the complex task could become as habitual as the simple task with more repetitions. Nevertheless, Wood et al. (2002) found that complex tasks were associated with more thoughts about the task during its performance than simple tasks, indicating they may be less automatic. It has also been suggested that the complexity of a task can determine the 'type' of automaticity which can develop. Wood and Neal (2007) have proposed that repetition of behaviour in a consistent setting could lead to the development of goal-directed automaticity rather than habit when the behaviour is complex. Because exercise behaviours can be considered more complex than eating or drinking behaviours, development of automaticity for these behaviours may take longer or reach a lower maximum. We therefore compared participants who had chosen these three different behaviours.

One important issue in understanding habit development is the role of omissions. James (1890) argued that development of habits requires uninterrupted performance. In the real world, this criterion is unlikely to be met since there would almost certainly be some occasions when the cue is encountered and the behaviour not performed (a missed opportunity), yet people still appear to acquire habits. One study explored the development of exercise habits over a 12-week period (Armitage, 2005) and found that in this context, lapses in performing the behaviour were negative predictors of future performance. However, behaviour was assessed in weekly blocks and a lapse was defined as not attending for a week. In this study, we consider how a single missed opportunity affects habit development. This is an important question for health interventions where the aim is to establish healthy habits; if missing a single opportunity to perform the behaviour is catastrophic, it would be very difficult to help people establish new healthy habits.

The present study therefore addressed the following questions: (i) Does an asymptotic curve model the relationship between repetition and subjective automaticity at an individual level, (ii) how many repetitions are needed to reach a plateau of automaticity, (iii) do time to asymptote or asymptote value vary by behaviour type and (iv) does missing an opportunity compromise habit development?

METHOD

Participants

The study was advertised to university students through an email circular. Interested participants ($n = 101$) attended an individual initial meeting, of whom 96 chose to participate in the study. A payment of £30 was offered for participating in the full 12 weeks of the study but was not contingent on behaviour change or habit development.

Procedures

At the initial meeting, participants were asked to choose a healthy eating, drinking or exercise behaviour that they would like to make into a habit. The behaviour had to be one that (i) they did not already do, (ii) could be performed in response to a salient daily event (cue) and (iii) had a cue that occurred every day and only once a day. Examples of the behaviours chosen were 'eating a piece of fruit with lunch', 'drinking a bottle of water with lunch' and 'running for 15 minutes before dinner'. Participants were asked to try to carry out the behaviour every day for 84 days. They were also asked to log on to the study website every day and report whether they had performed the behaviour the previous day as well as complete the SRHI for the behaviour. If they missed a report, they were permitted to retrospectively report whether they had performed the behaviour for up to three previous days, but the SRHI was only completed for the current day.

Measures

SRHI scores were the primary outcome measures (Verplanken & Orbell, 2003). The SRHI is a 12-item questionnaire answered on a 7-point Likert scale, with higher scores indicating stronger habit strength. The behaviour of interest is followed by statements to which participants report their level of agreement; example items are 'I do automatically', 'I do without thinking' and 'I would find hard not to do'. We created an automaticity subscale using 7 of the 12 items in the SRHI (items 2, 3, 5, 6, 8, 9, 10) and excluding the identity and repetition items, which gave a total score range of 0–42.

Analysis

The hypothesis in this study was that repeating a behaviour would result in increasing automaticity scores which would follow an asymptotic curve. To test this hypothesis, we plotted each individual's daily automaticity scores over 12 weeks to examine the shape of the individual automaticity curves, and examined the fit of an asymptotic curve. To identify communalities across individuals, we calculated the time taken for automaticity scores to plateau and the level of automaticity at the plateau. Ideally we would have looked at the relationship between number of repetitions and automaticity, but because not all participants logged on to the website every day, there were days for which behavioural data were unavailable. We therefore examined the relationship between day of the study and automaticity and considered consistency of performance as a separate predictor.

SPSS Version 14 was used to run nonlinear regressions to fit a curve for each individual's data based on Hull's (1943, 1951) model, using Mitscherlich's law of diminishing returns ($y = a - be^{-cx}$), where y is automaticity and x is day of the study (Mitscherlich, 1919). This equation has the advantage that the constants provide psychologically meaningful results: ' a ' represents the asymptote of the curve (the automaticity plateau score), ' b ' is the difference between the asymptote and the modelled initial value of y (when $x = 0$) and ' c ' is the rate constant that represents the rate at which the maximum is reached. All constants are positive numbers. The constants were used to calculate the time taken for the automaticity score to reach 95% of the asymptote using the following equation: Time to reach 95% of asymptote = $-\ln(a/20b)/c$. The R^2 value for the regression model was used to indicate how well the model fitted the data for each participant.

The SRHI does not have a pre-defined cut-off to identify a behaviour as habitual, but we took the decision to consider asymptote scores below 21 to indicate lack of habit. A score of 21 indicates an average response of 'neither agree nor disagree' for all the items in the subscale, which suggests that behaviours with SRHI scores lower than this cannot be considered a habit. This does not mean that scores of 21 and over are necessarily habitual, only that we were reasonably confident that scores below 21 are not.

Because participants were not entirely consistent in logging on to the website, there were days for which data were not available. Two variables were created to estimate the number of repetitions of the new behaviour during the study period. The first (number of reported repetitions) was calculated by simply summing all the occasions when a participant reported that they had performed the behaviour. The second (per cent compliance) was calculated as the percentage of all days for which data were reported that the participant reported having performed the behaviour.

In order to confirm that the modelled curve provided a better fit to an individual's data than a linear model, linear regressions were run for each individual's data, and R^2 values for the linear and nonlinear regressions compared using a sign test.

Many of the regression parameters and estimates of performance were not normally distributed. Therefore, unless otherwise stated, medians and quartiles ($Q_1:Q_3$) are presented and non-parametric tests are used to compare groups and examine associations between variables (Spearman's correlation).

RESULTS

The 96 participants (30 men, 66 women) were predominantly postgraduate students (two were undergraduates) with a mean age of 27 (range 21–45). The majority (65%) were White and born in the UK or Europe. For the habit task, 27 chose an eating behaviour, 31 a drinking behaviour (most often water), 34 an exercise behaviour and 4 another behaviour (e.g. meditation). The median automaticity score at baseline was 3 ($Q_1:Q_3 = 0:9$). Fourteen participants did not enter data

beyond day 60 and were considered to have dropped out of the study. The remaining 82 participants logged on an average (median) of 47 out of 84 days, and in combination with the occasions on which they entered behaviour data retrospectively, this gave an average of 76 days for which data were available. At the start of the study, Cronbach's α was .90 for the SRHI automaticity sub-scale.

Does an Asymptotic Curve Model the Relationship Between Repetition and Automaticity?

Individual scatterplots of automaticity scores over days were generated, which included information on whether or not the behaviour had been carried out by using different coloured markers (examples are shown in Figure 1). Visual inspection indicated that for many participants automaticity progressively increased during the study and the graphs showed the expected curve shape, with a steady increase over time until a final plateau. Therefore, it was appropriate to run nonlinear regressions fitting asymptotic curves.

When nonlinear regressions for curve-fitting were run for each participant, a variety of different results were obtained. For 12 participants, SPSS was unable to find an optimal solution after 100 iterations attempting to fit the curve equation to the data. For 8 participants, the model generated a zero value for 'b', indicating a horizontal line as the plotted shape for the data, which suggested that these participants did not increase their automaticity ratings over the study period. For the remaining 62 participants, SPSS was able to fit the model and the median R^2 value, for these participants, was .84 ($Q_1:Q_3 = 0.68:0.93$). The distribution of curve-fits was plotted and the threshold for 'adequate' fit was based on examining the distribution (Figure 2). A step-change was evident in the number of participants with R^2 values above 0.7 and so we used this as the threshold. For 16 participants, the R^2 value was below .7 which suggested a different pattern from the majority.

Four participants had an asymptote value of less than 21 and were therefore categorized as showing some increase in automaticity, but not to the point of having formed a habit. In other cases, the modelled values for constant 'a' exceeded the possible range on the automaticity scale because the model was not constrained by range. We accepted this as a feature of the model except where the modelled value of the asymptote became unrealistic. Given that the maximum (real) score for the automaticity total was 42, we allowed modelled asymptote values up to 49, which excluded three participants from the 'good fit' group. The model was therefore a good fit for 39 of the 82 participants (48%).

Table 1 shows the average values for the performance variables across the different categories of participants. Individuals whose curves had a good fit were relatively consistent in performing their chosen behaviour, as were those for whom the modelled asymptote was very high. Those for whom constant 'b' was modelled as 0, the asymptote was low, or the R^2 value was less than .7, were less consistent in performing the behaviour. The 12 participants for whom the model could not be fitted reported a moderately high level of performance, but examining these individuals' graphs visually, five appeared to change very little over the 84 days, and five looked as if they were increasing linearly over time.

Within the 82 participants with adequate data, the unadjusted R^2 from linear regressions was compared to the R^2 from the nonlinear regressions. Where a model could not be fitted these values were set as zero. The median linear R^2 was .66 ($Q_1:Q_3 = 0.40:0.81$) and the median nonlinear R^2 was 0.74 ($Q_1:Q_3 = 0.0:0.89$). A sign test showed this to be a significant difference ($Z = -3.019, p = 0.003$) with the nonlinear regressions performing better.

For those we classified as having a 'good fit' ($N = 39$) with the curve model, the median unadjusted R^2 from linear regressions was 0.72 ($Q_1:Q_3 = 0.55:0.80$) and the R^2 from the nonlinear regressions was 0.88 ($Q_1:Q_3 = 0.84:0.94$). A sign test showed this was also a significant difference ($Z = -5.765, p < 0.001$).

How Many Repetitions are Needed to Reach a Plateau of Automaticity?

Table 2 shows the values for the curve parameters for the 39 participants for whom the model was a good fit, and Figure 3 shows examples of the modelled curves of their data. Ten of these participants were performing eating behaviours, 15 drinking behaviours and 13 exercise behaviours. The median time to reach 95% of asymptote was 66 days, with a range from 18 to 254 days.

There was substantial variation in curve parameters across individuals. One explanation could be the number of repetitions of the behaviour. Among the 39 participants for whom the model was a good fit, correlations between

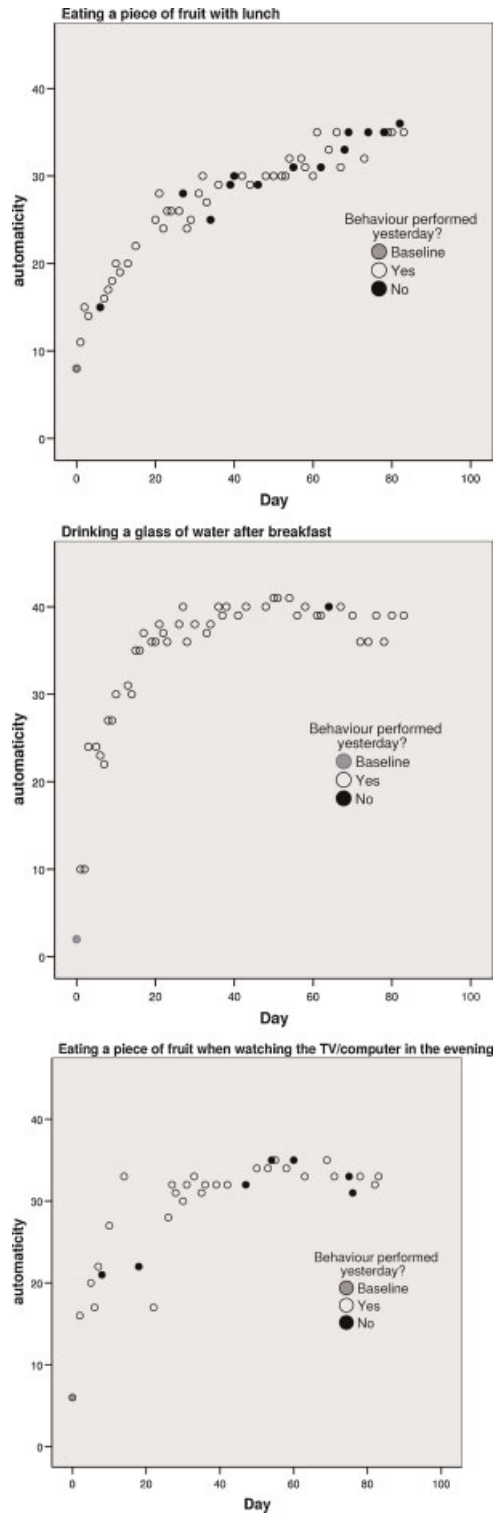


Figure 1. Three examples of increases in automaticity score during the 84 days of the study, showing on which days each participant had performed their chosen behaviour the previous day

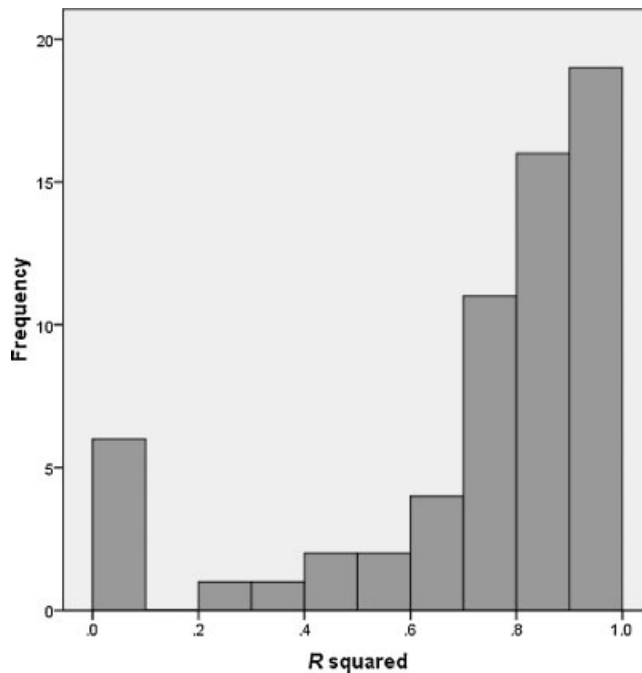


Figure 2. Histogram showing the distribution of R^2 values for the 62 participants the model was able to fit

Table 1. Performance variables within participant groups

	Good fit	$a > 49$	$a < 21$	$R^2 < .7$	b is modelled as 0	Model cannot fit
N	39	3	4	16	8	12
Percentage compliance, median (quartiles, $Q_1:Q_3$)	88 ^a (80:94)	82 ^{ab} (64:94)	61 ^b (37:73)	60 ^b (38:86)	63 ^b (48:73)	80 ^{ab} (52:95)
Number of reported repetitions, median (quartiles, $Q_1:Q_3$)	63 ^a (56:75)	52 ^{ab} (51:65)	44 ^b (27:52)	44 ^b (28:70)	40 ^b (35:51)	59 ^{ab} (36:72)

^{abc}Mann–Whitney tests between each pair of groups indicated significant differences between subgroups with different superscripts.

performance variables and the five curve parameters were calculated. The only significant correlation was between per cent compliance and nonlinear model R^2 ($r = .34, p = .035$), indicating that individuals who performed the behaviour more consistently showed a change in automaticity scores which was modelled more closely by an asymptotic curve. Time to reach 95% asymptote and level of asymptote (a) were also significantly correlated ($r = .38, p = .02$), indicating that reaching a higher asymptote took longer.

Table 2. Median, quartiles and minimum and maximum values of the curve parameters for those for whom the model was a good fit ($N = 39$)

	Median	Quartiles ($Q_1:Q_3$)	Minimum	Maximum
R^2	0.88	0.84:0.94	0.72	0.98
a (asymptote)	35	29:43	21	48
b (change in habit strength/automaticity)	30	23:39	13	51
c (rate of change)	0.042	0.028:0.064	0.010	0.170
Time to reach 95% of asymptote (days)	66	39:102	18	254

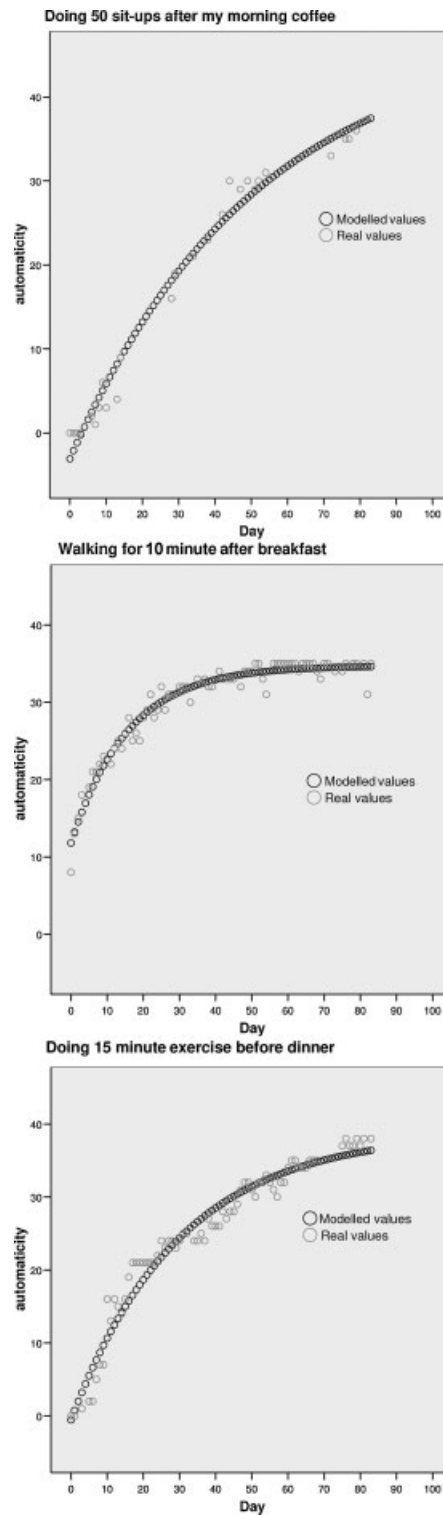


Figure 3. Three examples of increases in automaticity scores during the 84 days of the study showing the scores entered and the curve modelled using nonlinear regressions

Do Time to Asymptote or Asymptote Value Vary by Behaviour Type?

Among those with good model fit, the behaviours chosen can be categorised into eating ($N = 10$), drinking ($N = 15$) or exercise behaviours ($N = 13$). The median asymptote values were 34 ($Q_1:Q_3 = 27:44$), 35 ($Q_1:Q_3 = 29:43$) and 35 ($Q_1:Q_3 = 31:44$), respectively. A Kruskal–Wallis comparison showed that this was not a significant difference ($p = .708$). Times to reach 95% of the asymptote were 65 ($Q_1:Q_3 = 35:106$), 59 ($Q_1:Q_3 = 39:75$) and 91 ($Q_1:Q_3 = 44:118$) days respectively, which a Kruskal–Wallis comparison showed was also not a significant difference ($p = 0.328$).

Percentage compliance across the three behaviours was 80% ($Q_1:Q_3 = 77:90$), 93% ($Q_1:Q_3 = 86:96$) and 86% ($Q_1:Q_3 = 79:92$), respectively. A Kruskal–Wallis comparison showed that this was a significant difference ($p = 0.046$). Mann–Whitney tests showed the significant difference fell between the water drinking group (higher compliance) and the other two behaviour categories (lower compliance).

Does Missing an Opportunity Compromise Habit Development?

To test the hypothesis that missing an opportunity impacts development of automaticity, we compared automaticity scores on the days before and after a missed opportunity. For these analyses, a missed opportunity was defined as an occasion when the behaviour was reported not to have been performed but was immediately preceded by three occasions when it had been performed. One hundred and forty such missed opportunities were identified across 55 participants. Immediately following a missed opportunity, automaticity decreased by an average of 0.29 points on the SRHI automaticity subscale; a very small decrease. For some of these occasions ($N = 67$), we had SRHI scores for the subsequent day, after the behaviour had been performed again. The difference between the score before and after the omission was an increase of .55 points; which a Wilcoxon signed rank test showed was not significant.

In order to compare the effect of omissions with situations where the behaviour was performed consistently, we identified all the occasions when the behaviour was performed for three consecutive days and automaticity scores were available for the first and third days. Here the average increase across these three days was .79 points, comparable to the .55-point increase when there had been a missed opportunity. These results suggest that missing a day resulted in a non-significantly lower SRHI score the following day, but there were no longer-term costs associated with a single omission.

The missed opportunities analysis was repeated restricting it to participants for whom the model fitted well, but the results were similar. Nor was the effect related to the time in the habit formation process that the omission occurred (i.e. day of the study) ($r = 0.099$, $p = 0.246$, $N = 140$).

A Validity Check

There were no significant correlations between the total number of times each participant completed the questionnaire for their chosen behaviour and the curve parameters among the 39 for whom the model was a good fit. The median number of times the questionnaire was completed for these 39 participants was 50 ($Q_1:Q_3 = 35:64$).

DISCUSSION

This study used a novel approach to investigate habit formation in a real-world setting. For the majority of participants, automaticity increased steadily over the days of the study supporting the assumption that repeating a behaviour in a consistent setting increases automaticity. We provided no extrinsic rewards, indicating that they are not required for habit development, although because the behaviours were selected by the participants they were likely to have been intrinsically rewarding.

The aim of this study was to use an idiographic approach to model the habit formation process in individuals and assess whether an asymptotic model reflects a generalized habit formation process. An asymptotic model proved to be a good fit for almost half (48%) of the participants who provided enough data for analysis. Those for whom the asymptotic model was a poor fit had typically carried out the behaviour fewer times during the study. Two other groups of participants had relatively high levels of performance; one for whom the model could not be fitted and one with a very high modelled

asymptote. It is probable that these individuals were relatively slow in forming their habits and would have reached a plateau if the recording had continued for longer.

On average, across all participants with sufficient data (with those where the model was not able to fit assigned an R^2 value of 0), as well as among those with good curve fit, the fit of the asymptotic curve was superior to the linear model. We are therefore reasonably confident that the asymptotic curve reflects a generalized habit formation process. Early repetitions result in larger increases in automaticity than those later in the habit formation process, and there is a point at which the behaviour cannot become more automatic even with further repetition. This finding makes an important contribution to the literature because although this development process has been assumed in some habit models, it has not previously been tested for real world behaviours.

Modelling the habit formation process in individuals allowed us to calculate the time taken for automaticity scores to plateau; operationalized here as time to reach 95% of asymptote. The average modelled time to plateau in this sample was 66 days, but the range was from 18 to 254 days. We were only able to find one statement in the literature discussing how long it takes to form a habit. This was from Ronis, Yates, and Kirscht (1988), who argued that a behaviour is habitual once it has been 'performed frequently (at least twice a month) and extensively (at least 10 times)' (p. 213). Our study has shown that it is likely to take much longer than this for a repeated behaviour to reach its maximum level of automaticity. Interventions aiming to create habits may need to provide continued support to help individuals perform a behaviour for long enough for it to be subsequently enacted with a high level of automaticity. It is interesting to note that even in this study where the participants were motivated to create habits, approximately half did not perform the behaviour consistently enough to achieve habit status.

As discussed in the introduction, the habit idea was developed with simple behaviours, and there has been debate over whether complex behaviours can become automatic and whether the automaticity that develops for more complex tasks can be considered a habit. Among the successful habit-formers in the current study, those who chose drinking behaviours performed these more often than those who chose exercise behaviours, and these in turn were performed on more days than eating behaviours. There were no significant differences between groups on the curve parameters, but because the study was not powered for sub-group analyses, this negative finding should not be considered to be definitive. It is notable that the exercise group took one and a half times longer to reach their asymptote than the other two groups. Given that exercising can be considered more complex than eating or drinking, this supports the proposal that complexity of the behaviour impacts the development of automaticity.

The potential effect of errors of omission was highlighted by James (1890) but has not been tested formally before. We examined specific missed opportunities and found that they reduced the automaticity score by less than half a point. In instances where the behaviour was performed again the following day and automaticity data were available, scores increased from the day before the missed opportunity. Comparing increases in automaticity score over 2 days when the behaviour was either performed on both days (mean increase = 0.79) or there was a missed opportunity in between (mean increase = 0.55) shows that a missed opportunity did not materially affect the habit formation process. This is encouraging for researchers designing health interventions which aim to establish healthy habits. It suggests that although repetition of a behaviour is required in order to form a habit, some missed opportunities will not derail the process.

Armitage (2005) showed that lapses had important consequences for future performance. In his study, behaviour was assessed in weekly blocks and a lapse was defined as not attending for a week. This is markedly different from the single missed opportunities we have assessed. These findings can easily sit side-by-side and suggest that missing one opportunity does not preclude habit formation, but missing a week's worth of opportunities reduces the likelihood of future performance and hinders habit acquisition. Within the 39 participants in this study with a 'good fit' model, there was no relationship between performance estimates and plateau height. However, the four participants, for whom the model fitted well but had low automaticity scores, had performed the behaviour significantly less often than those with higher plateau scores. There may be a threshold value of performance below which maximum possible habit strength is curtailed. Above this threshold, it would appear that something other than mere number of repetitions determines the final habit strength.

We did not find a relationship between the timing of a missed opportunity and the change in habit score (i.e. missing an occasion early in the habit-formation process was no worse than missing an occasion later on), although there were relatively few such opportunities to evaluate and there is therefore a need for more work on this question.

There were a number of limitations to this study. The size of the sample of individuals with good curve fit was relatively small, so we had limited power to examine differences between participants. As expected, given the 3-month time frame, there were days when participants did not enter their behaviour data, which was a limitation in interpreting the results. The

SRHI is a self-report measure, and has never previously been used repeatedly as it was in this study. It is difficult to assess the extent to which completing the same questions every day affects people's responses, although the lack of correlation between the number of questionnaire completions and any curve parameter suggests that completing the questionnaire did not inflate reports of automaticity, and supports the use of this scale as a repeated measure.

There could be concern about ceiling effects in this study, i.e. that the plateau was an artefact of participants reaching the highest score on the scale. However, although the third quartile for the modelled asymptote was 43, these high asymptotes were often due to participants not reaching a plateau within the 84 days and the fitted curve estimating a high value for the asymptote. The actual highest scores reached varied greatly and only a small number of cases ($N = 7$) reached 42 (the highest possible score). Therefore plateaus were not due to ceiling effects.

Because this is a novel study design, decisions had to be made about how to examine the data. We choose to use the number of days it took participants to reach 95% of their asymptote of automaticity as an outcome. The essential nature of an asymptote—that it is never reached—means it is impossible to assess the time to reach 100% of asymptote, and anything above 95% leads to a very large estimate of time taken because of the shape of the curve. From visual inspection, the points identified using 95% are approximately where the curves appear to reach asymptote, whereas 90% identifies a level clearly below where the curves level off.

We acknowledge that the findings from this study rest on the assumption that the SRHI subscale can capture the extent to which an individual performs a behaviour automatically. Although this measure has been used by many researchers to measure habit strength (e.g. Brug et al., 2006; Verplanken & Melkevik, 2008; de Bruijn et al., 2007), future work will need to assess whether similar habit formation models are generated from more objective measures of automaticity. As discussed in the introduction, the SRHI is the only measure currently available which can be used in this repeated way.

The SRHI items used here assess automaticity rather than habit strength. As discussed above, Wood and Neal (2007) have suggested that repeating complex tasks may result in goal-directed automaticity, which is distinct from habit (it is flexible and remains tied to the original goal). Using the SRHI in this study means that we have assessed the development of automaticity for performing these behaviours rather than specifically habit. New measures will be needed to disentangle these two forms of automaticity.

Modelling habit formation as we have done in this study could be a useful tool for researchers because it provides parameters that are interpretable and comparable between individuals. For example, as we have shown, people doing different behaviours can be compared to assess differences in the habit formation process. Modelling habits in this way has the potential to add to findings from studies that assess habit strength at one time point because the asymptote value tells us the maximum level of automaticity for a specific behaviour in an individual. When habit strength is assessed only once, it is not possible to identify each individual's stage of the habit formation process.

There remain many questions which can now be addressed using this model. Building directly on this study the first of these are: What level of consistency of behaviour is required to form a habit, is the habit formation process different for complex behaviours and do missed opportunities early in the habit formation process hinder habit acquisition more than those later in the process?

In conclusion, repeating a behaviour in response to a cue appeared to be enough for many people to develop automaticity for that behaviour. Although consistency in repetition is required, the degree of consistency is not yet known. There was variation both in the maximum automaticity reached and the time taken to plateau. The model used in this study gave us a way of quantifying these parameters and therefore provides a tool for exploring factors that may explain this variation. The range of times to reach a plateau shows that it can take a large number of repetitions for an individual to reach their highest level of automaticity for some behaviours, and therefore creating new habits will require self-control to be maintained for a significant period before the desired behaviours acquire the necessary automaticity to be performed without self-control.

ACKNOWLEDGEMENTS

The work was carried out and supported by researchers at the Health Behaviour Research Centre which is funded through a Cancer Research UK Grant. The authors thank Mark Livermore for programming the study website and David Boniface for his suggestions on the analysis in this study.

REFERENCES

- Ajzen, I. (2002). Residual effects of past on later behavior: Habituation and reasoned action perspectives. *Personality and Social Psychology Review*, 6, 107–122.
- Armitage, C. J. (2005). Can the theory of planned behavior predict the maintenance of physical activity? *Health Psychology*, 24, 235–245.
- Bargh, J. A. (1994). The four horsemen of automaticity: awareness, intention, efficiency, and control in social cognition. In R.S. Wyer, & T. K. Srull (Eds.), *Handbook of social cognition: Vol 1 basic processes* (pp. 1–40). Hove: Lawrence Erlbaum Associates Publishers.
- Brug, J., de Vet, E., de Nooijer, J., & Verplanken, B. (2006). Predicting fruit consumption: Cognitions, intention and habits. *Journal of Nutrition Education and Behavior*, 38, 73–81.
- Chatzisarantis, N. L., & Hagger, M. S. (2007). Mindfulness and the intention-behavior relationship within the theory of planned behavior. *Personality and Social Psychology Bulletin*, 33, 663–676.
- de Bruijn, G., Kremers, S., de Vet, E., de Nooijer, J., van Mechelen, W., & Brug, J. (2007). Does habit strength moderate the intention-behaviour relationship in the Theory of Planned Behaviour? The case of fruit consumption. *Psychology and Health*, 22, 899–916.
- Hull, C. L. (1943). *Principles of behavior: An introduction to behavior theory*. New York: Appleton-Century-Crofts.
- Hull, C. L. (1951). *Essentials of behavior*. Westport, CT: Greenwood Press.
- James, W. (1890). *The principles of psychology*. London: Macmillan.
- Kremers, S. P., & Brug, J. (2008). Habit strength of physical activity and sedentary behavior among children and adolescents. *Pediatric Exercise Science*, 20, 5–14.
- Kremers, S. P., van der Horst, K., & Brug, J. (2007). Adolescent screen-viewing behaviour is associated with consumption of sugar-sweetened beverages: the role of habit strength and perceived parental norms. *Appetite*, 48, 345–350.
- McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective retrieval: A multiprocess framework. *Applied Cognitive Psychology*, 14, S127–S144.
- Mitscherlich, E. A. (1919). Zum Gesetz des Pflanzenwachstums. *Fuhlings landwirtschaftliche Zeitung*, 21–22, 419–426.
- Reinaerts, E., de Nooijer, J., Candel, M., & de Vries, N. (2007). Explaining school children's fruit and vegetable consumption: The contributions of availability, accessibility, exposure, parental consumption and habit in addition to psychosocial factors. *Appetite*, 48, 248–258.
- Ronis, D. L., Yates, J. F., & Kirscht, J. P. (1988). Attitudes, decisions, and habits as determinants of repeated behavior. In A.R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 213–239). Hillsdale: Lawrence Erlbaum Associates.
- Sheeran, P. (2002). Intention-behavior relations: A conceptual and empirical review. In W. Stroebe, & M. Hewstone (Eds.), *European review of social psychology* (Vol. 12, pp. 1–30). New York: Wiley.
- Verplanken, B., & Melkevik, O. (2008). Predicting habit: The case of physical exercise. *Psychology of Sport and Exercise*, 9, 15–26.
- Verplanken, B. (2006). Beyond frequency: Habit as a mental construct. *British Journal of Social Psychology*, 45, 639–656.
- Verplanken, B., Myrbakk, V., & Rudi, E. (2005). The measurement of habit. In T. Betsch, & S. Haberstroh (Eds.), *The routines of decision making* (pp. 231–247). London: Erlbaum.
- Verplanken, B., & Orbell, S. (2003). Reflections on past behavior: A self-report index of habit strength. *Journal of Applied Social Psychology*, 33, 1313–1330.
- Verplanken, B., Aarts, H., van Knippenberg, A., & van Knippenberg, C. (1994). Attitude versus general habit: Antecedents of travel mode choice. *Journal of Applied Social Psychology*, 24, 285–300.
- Wood, W., Quinn, J. M., & Kashy, D. A. (2002). Habits in everyday life: Thought, emotion, and action. *Journal of Personality and Social Psychology*, 83, 1281–1297.
- Wood, W., Tam, L., & Witt, M. G. (2005). Changing circumstances, disrupting habits. *Journal of Personality and Social Psychology*, 88, 918–933.
- Wood, W., & Neal, D. T. (2007). A new look at habits and the habit-goal interface. *Psychological Review*, 114, 843–863.