Reflections on Past Behavior: A Self-Report Index of Habit Strength[1]

BAS VERPLANKEN[2]
University of Tromsø
Tromsø, Norway

SHEINA ORBELL
University of Essex
Essex, United Kingdom

We argue that habit is a psychological construct, rather than simply past behavioral frequency. In 4 studies, a 12-item index of habit strength (the Self-Report Habit Index, SRHI) was developed on the basis of features of habit; that is, a history of repetition, automaticity (lack of control and awareness, efficiency), and expressing identity. High internal and test–retest reliabilities were found. The SRHI correlated strongly with past behavioral frequency and the response frequency measure of habit (Verplanken, Aarts, van Knippenberg, & van Knippenberg, 1994). The index discriminated between behaviors varying in frequency, and also between daily vs. weekly habits. The SRHI may be useful as a dependent variable, or to determine or monitor habit strength without measuring behavioral frequency.

When was the last time you performed a new behavior? In everyday life, it does not occur so often that we do something for the first time: Repetition is the rule, rather than the exception. The role of previous behavior has long been an issue of debate in the area of attitude–behavior relations (e.g., Eagly & Chaiken, 1993). A consistent finding is that measures of past behavioral frequency predict the occurrence of future behavior over and above established antecedents of behavior such as attitudes and intentions (e.g., Ouellette & Wood, 1998; Sutton, 1994). However, the debate about the role of past behavior has stalled for a long time on the conclusion that a statistical relation between the frequency of past and later behavior is not insightful by itself because frequency of past behavior is a construct that has no explanatory value (e.g., Ajzen, 1991; Ajzen & Fishbein, 2000; Eagly & Chaiken, 1993).

Residual shared variance between measures of frequency of past and later behavior (e.g., after attitudes and intentions have been taken into account) might reflect the influence of unmeasured variables on both occasions. This indeed is

[1]The authors thank Rob Holland and Raymond Smeets for their contribution in collecting the data of Experiment 1, and Kenneth Lindberget for collecting the data of Experiment 2. They also thank Icek Ajzen, David Silvera, and two anonymous reviewers for their useful comments on an earlier version of this article.

[2]Correspondence concerning this article should be addressed to Bas Verplanken, Department of Psychology, University of Tromsø, N-9037 Tromsø, Norway. E-mail: verplanken@psyk.uit.no

Journal of Applied Social Psychology, 2003, **33**, 6, pp. 1313-1330.
Copyright © 2003 by V. H. Winston & Son, Inc. All rights reserved.

not a particularly revealing conclusion. A perhaps significant twist in this debate is the acknowledgment that repeated behavior might become habitual (cf. Ajzen, 2001; Ronis, Yates, & Kirscht, 1989; Triandis, 1977; Verplanken & Aarts, 1999). By satisfactory repetition, a behavior might become automatic in the sense that a specific response is spontaneously triggered by a specific cue in the environment. A statistical relationship between past and future behavior thus might become more meaningful. Once behavior has been sufficiently repeated, it might be habit—that is, the automaticity in responding to certain cues, and not frequency of behavior as such, that determines the occurrence of future behavior. Therefore, if one acknowledges that repeatedly performed behaviors might gain habitual qualities, there is good reason to develop an instrument that measures the degree to which a behavior has acquired that quality. In other words, if habit is a psychological construct, and not simply past behavioral frequency, it should be measured as such.

Habits can be defined as "learned sequences of acts that have become automatic responses to specific cues, and are functional in obtaining certain goals or end-states" (Verplanken & Aarts, 1999, p. 104). Thus, the first feature of habits is that they have a history of repetition. The more frequently we perform a behavior, the more likely it becomes habitual.

However, it is not the recurrence of a behavior per se that constitutes a habit. Rather, a habit is created by frequently and satisfactorily pairing the execution of an act in response to a specific cue (e.g., Hull, 1943; Tolman, 1932). This process results in a mental representation of an association between a goal and an action. For example, Aarts and Dijksterhuis (2000) measured bicycle use habit and manipulated the activation of travel goals (e.g., going shopping). In a subsequent task, they measured the time it took to indicate whether using the bicycle was suitable for specific travel locations (e.g., a particular shopping mall). They found that habitual bicycle users showed faster responses than did nonhabitual bicycle users, but only if travel goals had been primed.

These results demonstrate that when a goal that is associated with a habit is activated, specific cue–response links become more accessible. Encountering such a cue thus automatically triggers the habitual response. In other words, whereas new behavior may follow from conscious decision making, the formation of a habit implies the delegation of control over the behavior to the environment (e.g., Bargh & Gollwitzer, 1994; Verplanken & Aarts, 1999). For example, you might discover that taking the 8:10 a.m. bus gives you the best connection to your new job. Once you have satisfactorily practiced this connection, an association is created between a particular cue (8:00 in the morning) and a particular behavior (going to the bus stop). If the goal of going to work is activated, encountering that cue will automatically trigger the behavior. No conscious decision making is needed anymore: A habit has been born. Note that goal activation is an important element. You do not go to the bus stop at 8:00 on Sunday mornings. Thus,

although a history of repetition remains a fundamental characteristic of a habit, the habit construct includes more than that; in particular the aspect of automaticity.

Measurement of Habit Strength

Because habits develop and gain strength by satisfactory repetition of behavior (e.g., Hull, 1943; James, 1890; Tolman, 1932; Triandis, 1977, 1980; Watson, 1914), it is not unreasonable to accept a measure of past behavioral frequency as a measure of habit strength. Such a measure is used commonly by social psychologists (e.g., Eagly & Chaiken, 1993; Ouellette & Wood, 1998; Verplanken & Aarts, 1999). Some researchers use observation or objective registration of behavior to obtain an estimate of behavioral frequency (e.g., Fredricks & Dossett, 1993; Landis, Triandis, & Adamoupoulos, 1978). In most cases, self-reports of past behavioral frequency are used; for instance, "How many times during the past 2 weeks did you use your seat belt?" However, if one acknowledges that habit is a psychological construct, rather than past behavioral frequency, such measures do not seem adequate.

There are reasons why one might wish to have a measure of habit strength that directly taps the construct (in particular the automaticity aspect), rather than merely an estimate of behavioral frequency. For example, one might want to test whether an intervention increases the frequency of a particular behavior and makes it become habitual (e.g., dental flossing) or decreases the frequency and loses its habitual quality (e.g., choosing the car vs. public transportation). In such cases, pre- and post-measures of habit strength are required, and one might prefer measures that directly tap the habit construct to a mere estimate of behavioral frequency. An independent measure of habit strength is also required if one wishes to determine the effect of habit in addition to behavioral frequency (cf. Ajzen & Fishbein, 2000). Finally, there might be contexts in which behavioral frequency measures have the danger of being unreliable (e.g., Schwarz, 1999); for instance, when it is difficult to retrieve episodic memories (e.g., it is hard to remember how many times during the last month one took the car) or when behavior itself is fuzzy (e.g., eating unhealthy food).

Which alternative measures of habit are available? Verplanken, Aarts, van Knippenberg, and van Knippenberg (1994) introduced a measure of habit that was based on the idea that a habit is a form of scripted behavior. In this case, scripts refer to idiosyncratic cognitive structures that represent associations between behaviors and goals in specific contexts (cf. Aarts & Dijksterhuis, 2000; Schank & Abelson, 1977; Trafimow & Wyer, 1993; Wittenbraker, Gibbs, & Kahle, 1983). This measure, which was later referred to as the *response-frequency measure of habit* (Verplanken & Aarts, 1999), presents participants with a number of habit-related situations (e.g., travel destinations) and asks them to respond as quickly as possible to generate the behavioral choice option that

they associate with that situation (e.g., taking the car). The assumption is that the more habitual that response is, the more frequently it is chosen across the set of situations. The proportion of these responses thus serves as a measure of habit strength (Aarts, Verplanken, & van Knippenberg, 1997, 1998; Verplanken, Aarts, & van Knippenberg, 1997; Verplanken, Aarts, van Knippenberg, & Moonen, 1998; Verplanken et al., 1994).

Although the response-frequency measure does not ask for behavioral frequency in the past, the measure cannot always be used and is not easy to administer. The use of the response-frequency measure is restricted to choice behavior that is executed in various contexts (e.g., using the car to go shopping, visit friends, go long or short distances, drive for pleasure). The measure also requires participants to respond as quickly as possible, which requires a controlled research environment (e.g., a research assistant to impose time pressure on the participants). The response-frequency measure also requires extensive pilot work for each new behavior so as to find an appropriate set of situations that constitute the measure. In other words, although the response-frequency measure has been found to be a satisfactory alternative to measures of past behavioral frequency in some contexts, there are a number of limitations to its use.

Another approach to a measure of habit is to ask participants directly to report on habit strength. For example, some researchers asked how often one conducted a particular behavior in the past "without awareness" or "by force of habit" (e.g., Kahle & Beatty, 1987; Mittal, 1988; Towler & Shepherd, 1991; Wittenbraker et al., 1983). These self-report measures have two serious problems. The first is that participants are asked simultaneously to provide both an estimate of behavioral frequency and an indication of the degree to which behavior is habitual in one response. Behavioral frequency and habit strength are thus confounded in this measure. Second, single-item measures that are supposed to tap a latent construct (e.g., habit strength) are notoriously unreliable and imprecise (e.g., Spector, 1992). However, although the self-reported habit strength measures used so far are problematic, there is no reason why habit strength might *not* be measured by means of self-reports. We think that, if measured appropriately, it is possible to have people reflect on their behavior in terms of the degree to which it is habitual.

At least two conditions for a well-designed measure should be fulfilled: The instrument should have a theoretically sound foundation, and it should be a multiple-item instrument. A theory-based measure is preferable over a measure based on face value so as to be more confident that the measure is valid. Multiple items are needed for reasons of reliability. Multiple items also are required when the measured construct is multifaceted, which is the case with habit. In the following section, we describe the theoretical rationale for a self-report index of habit strength and present a 12-item instrument. We then report four studies that provide a first test of the reliability and validity of the instrument.

A Self-Report Index of Habit Strength

The definition of habit that we cited at the beginning of the previous section highlights a number of features. The first is that habits have a history of repetition. The more frequently we perform a behavior, the more likely it is to be habitual. Importantly, as was argued earlier, a habit is a form of automaticity; in particular, the automatic elicitation of behavior upon encountering specific cues in the context of an activated goal.

Bargh (1994, 1996) distinguished four distinct and independent features by which an automatic process might be characterized: unintentionality, uncontrollability, lack of awareness, and efficiency. A particular process might be characterized by each of these four features being present or absent, which thus may lead to a number of variants of automaticity. Habit can be characterized as behavior that is intentional in its origin, is controllable to a limited extent, is executed without awareness, and is efficient.

Most habits that are studied by social psychologists are intentional. Intentionality in this case does not refer to planned or conscious action, but rather to being functional and goal directed. Habits are developed as behaviors that appear to be satisfactory in fulfilling some goal (e.g., taking the car to reach a destination efficiently, using seatbelts for safety, brushing teeth to keep ourselves healthy, eating fatty food for pleasure). Although most habits are, in principle, controllable (e.g., by deliberate thinking and planning), it often appears difficult to overrule strong habits (e.g., Aarts & Dijksterhuis, 2000; Heckhausen & Beckmann, 1990; Verplanken & Faes, 1999). Finally, habits are efficient in the sense that they free mental capacity to do other things at the same time. The efficiency of habits appears in particular under conditions of heavy load, such as exhaustion, time pressure, distraction, or information overload.

We assumed that uncontrollability, lack of awareness, and efficiency would be the features that characterize the experience of habits in everyday life. In addition to a history of repetition and features of automaticity, we added another aspect of habits; namely, the fact that habits are part of how we organize everyday life and thus might reflect a sense of identity or personal style. We thus assumed that although this might not hold for all habits, at least some might be descriptive of a person and thus express someone's identity (cf. Trafimow & Wyer, 1993).

We designed a self-report instrument to measure habit strength that was based on the features discussed earlier; that is, the history of repetition of behavior, the difficulty of controlling behavior, the lack of awareness, efficiency, and the identity element. The instrument contains 12 items and will be referred to as the Self-Report Habit Index (SRHI; Appendix A). Four studies were conducted to provide indications of reliability and validity of the SRHI. Internal structure and reliability of the SRHI were investigated by principal components analyses and

coefficient alphas in all studies. Study 1 inspected the test–retest reliability of the index. In Study 2, convergent validity was tested by relating the SRHI to the response-frequency measure (Verplanken et al., 1994) as an alternative independent measure that taps the automatic qualities of habitual behavior. Study 3 was conducted to provide further convergent validity by investigating correlations between the SRHI and behavioral frequency measures with respect to three behaviors that varied in frequency of occurrence. Finally, Study 4 tested whether the SRHI was able to discriminate between weekly and daily personal habits.

Study 1

Method

Participants were 93 undergraduate students (61 females, 32 males) at the University of Nijmegen, The Netherlands. This study was incorporated as a filler task in an unrelated experiment. Participants came twice to the laboratory with a 1-week delay between the two sessions. They were told that the experiment was on travel-mode choices and responded to the SRHI concerning using the bicycle in town. Responses were made on a 7-point Likert scale ranging from 1 (*agree*) to 7 (*disagree*). The scores were recoded such that high values indicated strong habits. Most participants (N = 86, 92%) returned 1 week later. In the second session, the same SRHI was again incorporated as a filler task.

Results and Discussion

Principal components analyses were conducted on the 12 items of the SRHI on both occasions. On both occasions, three eigenvalues greater than 1 appeared: 5.73, 1.34, and 1.15 for the pretest index; and 6.58, 1.55, and 1.05 for the posttest index. The pattern of the eigenvalues clearly suggests a one-dimensional structure. The first component accounts for 47.76% and 54.84% of the variance for the pretest and posttest indexes, respectively. Coefficient alphas of the pretest and posttest indexes are .89 and .92, respectively, which indicate high internal reliabilities.

The items of each respective SRHI were averaged. The mean SRHI scores in the pretest and posttest were 4.84 (SD = 1.23) and 4.73 (SD = 1.33), respectively. A correlation of .91 ($p < .001$) between the pretest and posttest indexes indicates a high test–retest reliability.

Study 2

In this study, the convergent validity of the SRHI was tested by relating it to the response-frequency measure of habit (Verplanken et al., 1994). The

response-frequency measure is supposed to tap the degree to which choices are scripted and appear across a number of situations. This measure is thus considered an alternative measure of the automatic qualities of habitual behavior.

Method

Participants wre 86 (51 females, 35 males) undergraduate students at the University of Tromsø, students at a school for physiotherapy, and students at a high school for fishery in Tromsø, Norway, respectively. The experiment focused on transportation mode choices, and the target behavior was taking the bus.

A response-frequency measure of bus habit was composed of 12 travel destinations (Verplanken et al., 1994). The destinations were located inside the town of Tromsø (e.g., "a bar in the town center")—within 7 km from the town center (e.g., "Tromsø Museum"); or further away than 7 km, but reachable by bus (e.g., "a friend in Kvaløysletta"). Participants were asked to choose as quickly as possible among five modes of transportation (i.e., car, walking, bus, bicycle, or cab, respectively). The number of times the bus was chosen served as a measure of habit strength. The instrument was taken as a filler task in an unrelated experiment. Later in the (unrelated) experiment, the SRHI was presented as another filler task. Participants responded to the stem "Taking the bus is something . . .," and were presented with the 12 SRHI items. Responses were made on a 7-point Likert scale ranging from 1 (*agree*) to 7 (*disagree*). The scores were recoded such that high values indicate strong habits.

Results and Discussion

The response-frequency measure of habit strength scores concerning taking the bus ranged from 0 to 12 ($M = 5.19$, $SD = 3.94$). The SRHI items were subjected to a principal components analysis. Three eigenvalues were greater than 1 (5.80, 1.23, and 1.12, respectively), while the first component accounted for 47.32% of the variance. The pattern of eigenvalues suggests a one-dimensional structure. Coefficient alpha of the SRHI was .89. The scale mean was 4.77 ($SD = 1.34$). The SRHI correlated strongly and significantly with the response-frequency measure ($r = .58$, $p < .001$). The size of this correlation thus testifies to the convergent validity of the SRHI.

Study 3

The goal of this study was to establish convergent validity of the SRHI by correlating the index with estimates of past behavioral frequency of three behaviors that differed substantially in frequency. The study thus also tested whether the SRHI was able to differentiate between these behaviors in terms of habit strength.

Method

Participants were 143 undergraduate students (100 females, 43 males) at the University of Nijmegen, The Netherlands, who were paid for their participation. Participants came to the laboratory twice with a 3-week delay between the two sessions. Of these participants, 10 (7%) did not show up at the second session, thus leaving 133 participants in the study.

During the first session, participants were presented with 26 behaviors and were asked to indicate the frequency with which they engaged in these behaviors. For each behavior, participants indicated frequency by selecting one of the following choices: *never*, *once a month*, *twice a month*, *three times a month*, *once a week*, *twice a week*, *three times a week*, *four times a week*, *five times a week*, *six times a week*, *once a day*, *twice a day*, *three times a day*, *four times a day*, *five times a day*, or *more than five times a day.* These responses were coded from 0 to 15 respectively (low values indicate low frequency).

From the 26 behaviors, 3 were selected to be included in the second session 3 weeks later, such that the behaviors varied in reported frequency in the first session. The three behaviors were "Watching *Goede Tijden, Slechte Tijden* [Good Times, Bad Times] (GTST)," which was a well-known soap on Dutch television at that time (broadcast five times a week); "Eating candies"; and "turning on music at home." The frequency measures of the three behaviors indicated that they were executed, on average, about three times a month, four to five times a week, and twice a day, respectively. During the second session 3 weeks later, the SRHI was administered for each of the three behaviors during a computer-controlled session. Responses were provided on an 11-point Likert scale ranging from 1 (*agree*) to 11 (*disagree*). Items were recoded such that high values indicate strong habits. The order of behaviors was random for each participant.

Results and Discussion

Principal components analyses were conducted on the SRHI for each behavior. In all three cases, a one-dimensional structure appeared. For watching GTST, two eigenvalues were greater than 1 (7.56 and 1.65), while the first component accounted for 62.98% of the variance. For eating candies, only the first eigenvalue was greater than 1 (6.58), and the first component accounted for 65.56% of the variance. For turning on music, two eigenvalues were greater than 1 (7.41 and 1.01), and the first component accounted for 61.73% of the variance. Coefficient alphas of the SRHI were high for each behavior: .94 for watching GTST, .95 for eating candies, and .94 for turning on music at home.

Table 1 shows the means and standard deviations of the behavioral frequency measures and of the SRHI, together with the correlations between the two measures for each of the three behaviors. As intended, the three behaviors differed significantly in frequency, as measured during the first session. This was

Table 1

Means and Intercorrelations of Behavioral Frequency Measures and Self-Reported Habit for the Three Behaviors (Study 3)

Behavior	Behavioral frequency		SRHI		Behavioral frequency: SRHI (*R*)
	M	*SD*	*M*	*SD*	
Watching GTST	4.07	3.45	3.47	2.41	.74
Eating candies	8.71	3.15	5.31	2.61	.55
Switching on music	11.85	2.46	7.18	2.45	.65

Note. SRHI = Self-Report Habit Index. GTST = *Goede Tijden, Slechte Tijden* [Good Times, Bad Times]. The measures of behavioral frequency ranged from 0 (*low*) to 15 (*high*). Self-reported habit ranged from 1 (*weak habit*) to 11 (*strong habit*). All correlations were significant at $p < .001$.

confirmed by a multivariate test on the three frequency means, $F(2, 136) = 227.71$, $p < .001$, and a highly significant linear trend that describes the mean frequencies of watching GTST, eating candies, and turning on music, respectively, $F(1, 137) = 456.57$, $p < .001$. Paired-sample t tests confirmed that the three behaviors differed significantly from each other, $t(138) = 11.96$, $p < .001$, and $t(141) = 9.26$, $p < .001$, for watching GTST versus eating candies, and eating candies versus turning on music, respectively.

The results of Study 3 provide two important findings. The first is that the SRHI correlated substantially with the behavioral frequency measure, which was taken 3 weeks earlier. The fact that the three behaviors were among a set of 26 behaviors during the first session and the relatively long delay of 3 weeks between sessions made it very unlikely that these correlations were inflated by a tendency to appear consistent. The second finding is that, as expected, the SRHI differentiated between behaviors that were assumed to differ in habit strength. This is indicated by a multivariate test on the three SRHI means, $F(2, 132) = 71.21$, $p < .001$; a highly significant linear trend that described the mean SRHI of watching GTST, eating candies, and switching on music, respectively, $F(1, 133) = 142.41$, $p < .001$; and paired-sample t tests, $t(133) = 6.32$, $p < .001$, and $t(133) = 6.10$, $p < .001$, for watching GTST versus eating candies, and eating candies versus switching on music, respectively. Thus, the level of habit strength as measured by the SRHI increased as a function of the level of behavioral frequency. Of course, because in this study the level of habit strength was confounded with type of behavior, the conclusion that the SRHI differentiates between levels of habit strength is not unequivocal. Therefore, we conducted a

fourth study, which was aimed at testing this hypothesis in a more unequivocal fashion.

Study 4

This study was aimed at testing the generality of the finding in Study 3 that the SRHI can differentiate between levels of habit strength. In particular, we tested whether habits that are performed daily versus habits that are performed weekly differ on the SRHI. The hypothesis was tested across a large number of different habits (a unique habit for each participant).

Method

Participants were 76 undergraduate students (52 females, 24 males) at the University of Nijmegen, The Netherlands, who were paid for their participation. As in the previous experiment, participants came twice to the laboratory, in this case with a 1-week delay between the two sessions; 2 participants (3%) did not show up at the second session, thus leaving 74 participants in the study.

In the first session, participants were asked to list two categories of their own habits: habits that are executed on a daily basis, and habits that are executed on a weekly basis. Participants also indicated how frequently they performed each habit (i.e., how many times per day in the daily category, and how many times per week in the weekly category). After this session, participants were randomly assigned to either a daily habit condition or a weekly habit condition (*n*s were 36 and 38, respectively).

For the second session, we selected for each participant a habit that was mentioned in the first session from the category to which the participant was assigned (i.e., either a daily or a weekly performed habit). Two criteria for the selection of a habit for a particular participant were used. First, the habit with the highest frequency within the participant's list in the assigned category was selected. Second, we wished to include as many different habits as possible, and preferably a unique habit for each participant. However, because some participants listed only one habit in the assigned category, some habits appeared more than once across the sample (i.e., 5 of the 31 daily habits and 5 of the 31 weekly habits appeared twice, and 1 of the weekly habits appeared three times). The result was that in both conditions, 31 different habits were presented. Appendix B displays the selected habits for each condition. During the second session, an SRHI with an 11-point response scale was presented for the selected habit.

Results and Discussion

First, we inspected the frequencies of the selected habits in order to be sure that the daily habits were indeed performed more frequently than were the

weekly habits (in theory, participants could have listed habits in the weekly category that were executed daily). The average frequency of the selected habit was 3.86 in the daily habits category (i.e., how many times the habit was executed every day) and 3.07 in the weekly habits category (i.e., how many times the habit was executed every week). If the daily habits are extrapolated to weekly frequencies, the frequencies in the daily habits condition were statistically significantly higher than those in the weekly condition, as expected (daily, $M = 27.02$; weekly, $M = 3.07$), $t(72) = 4.81$, $p < .001$. Even if one assumes that the daily behaviors were performed only 5 days a week (e.g., some behaviors are typically performed on working days or school days), the frequencies of daily and weekly behaviors differ highly significantly (daily, $M = 19.30$; weekly, $M = 3.07$), $t(72) = 4.55$, $p < .001$.

The SRHI (which in this case refers to a large variety of different personal habits) was subjected to a principal components analysis. Three eigenvalues were greater than 1 (4.62, 2.02, and 1.16), while the first component accounted for 38.48% of the variance. Although a one-dimensional structure is less obvious than in the previous three studies, we opted for that solution. This results in a scale with a satisfactory coefficient alpha of .85.

The main analysis consists of a comparison of the SRHI between the daily and weekly habit conditions. These differed significantly in the expected direction (daily habit, $M = 7.83$; weekly habit, $M = 6.82$), $t(72) = 2.31$, $p < .03$. The size of the means (given the 11-point scale) suggests that both categories represent behaviors that have habitual qualities. Nevertheless, and in spite of the small differences, the SRHI appears sufficiently sensitive to differentiate between daily and weekly habits. Thus, across a large number of personal habits, the measure indicates that habits that are performed three times daily are stronger than are habits that are performed three times weekly.

General Discussion

In four studies, a 12-item self-report instrument of habit strength (the Self-Report Habit Index) was tested for reliability and validity. The SRHI was developed as a direct measure of habit strength that does not rely on estimates of behavioral frequency. The SRHI showed high internal reliabilities in all studies.

Study 1 demonstrated high test–retest reliability between two measurements separated by a week. In Study 2, the SRHI correlated strongly and significantly with the response-frequency measure, which is another measure of habit strength focusing on the automatic qualities of habitual responses (Verplanken et al., 1994). The response-frequency measure is especially applicable to choices that are made in various contexts, such as travel mode choices, in this case. The SRHI is also appropriate for such contexts, but is more general in its applications as well.

In Study 3, the SRHI correlated consistently and significantly with behavioral frequency measures. Note that because behavioral frequency measures and the SRHI are very different instruments, these correlations are unlikely to be inflated by common measurement error. In addition, the SRHI discriminated between three behaviors that differed in average behavioral frequency, and thus should be expected to differ in average habit strength.

In Study 4, the findings that the SRHI discriminates between behaviors that differ in frequency was replicated across a large set of personal habits in a design in which behavioral frequency and type of behavior were not confounded. In a between-participants design, it was demonstrated that the SRHI was higher for 31 habits that are executed daily than for 31 habits that are executed weekly.

Taken together, the results suggest that the SRHI is a reliable and valid instrument. It measures habit strength by breaking it down into a number of features of habit; that is, history of repetition, automaticity (lack of control, lack of awareness, efficiency), and expression of one's identity. Including these features in an instrument undoubtedly provides a more complete and full account of habit, compared to existing instruments, particularly behavioral frequency measures.

Because the SRHI includes three items that refer to past behavioral frequency, one might argue that the results are not particularly new or might even be circular because we used measures of behavioral frequency as a criterion in Studies 3 and 4. First of all, of course, the SRHI items tap the experience of repetition, and thus differ from the frequency estimates that are used to validate the SRHI. However, in order to rule out any suspicion of circularity in our results, we reanalyzed the data in the four studies while leaving out the three frequency-related items (i.e., Items 1, 7, and 12) from all SRHIs. Coefficient alphas of the nine-item SRHIs remained very high in Studies 1, 3, and 4, and were acceptable in Study 2. Importantly, the main results in all studies remained unchanged, thus strongly suggesting that these are not solely dependent on the inclusion of the frequency-related items.[3]

How happy should we be with a new self-report instrument of habit strength? First and foremost, it is important to consider habit as a psychological construct that has a number of facets, rather than simply defining habit as past behavioral frequency. The SRHI was thus developed to measure this construct. Furthermore, there are many contexts in which we might wish to have an instrument that

[3]The exact results of the analyses with the three frequency-related items removed from the SRHIs were as follows. Coefficient alphas were .88 and .91 for the pre- and post-measures in Study 1, respectively; .61 in Study 2; .92, .93, and .93 for watching GTST, eating candies, and turning on music in Study 3; and .82 in Study 4. The test–retest correlation in Study 1 was .90 ($p < .001$); the SRHI/response-frequency measure correlation in Study 2 was .43 ($p < .001$); and the SRHI/behavioral-frequency correlations were .69, .54, and .63 (all ps $< .001$) for watching GTST, eating candies, and turning on music in Study 3, respectively. The difference in SRHI between the daily and weekly behaviors condition in Study 4 was significant, $t(72) = 2.31$, $p < .03$.

measures habit strength directly by addressing components of the habit construct, rather than an estimate of past behavioral frequency. For instance, the SRHI might be useful to establish whether a particular behavior gains or loses habitual qualities. Whereas interventions usually aim at forming new behavior or changing old behavior (e.g., promoting a healthier diet), there is too little attention to the habituation of the new behavior as a goal of an intervention (Verplanken & Aarts, 1999). Habits also may become part of implementation strategies (e.g., Orbell, Hodgkins, & Sheeran, 1997). In all these cases, a behavioral frequency measure is inadequate as a measure of habit, and the SRHI might be a useful alternative.

It might be useful to have an instrument of habit strength that is not based on estimates of behavioral frequency when we wish to establish the contribution of habit in addition to behavioral frequency. Thus, the SRHI might prove useful in discussion of the residual effect of past behavior in the prediction of future behavior (e.g., Ajzen, 2002). In any case, a theory-based multi-item instrument like the SRHI seems to be a great improvement over the existing single-item self-report measures of habit strength.

On the other hand, one might be suspicious of a self-report instrument that attempts to tap qualities like the extent to which a behavior is automatic (cf. Nisbett & Wilson, 1977). Thus, are people able to reflect on how habitual a particular behavior is? Asking such a question directly (e.g., "To what extent is Behavior X a habit?") would likely yield responses that lack validity and reliability. We therefore broke the concept of habit down into components that seem relatively easy to reflect on, such as the fact that habitual behavior is repetitive, difficult to control, goes with a lack of awareness, is efficient and may reflect one's identity. In our view, this might provide a way to obtain a valid and reliable indication of habit strength. From a practical point of view, the SRHI is an easy instrument, both for participants to respond to and for researchers to use. It can easily be included in questionnaires and does not require pilot studies, additional instructions, equipment, or a highly controlled research environment.

Some words of caution must be added. Undoubtedly, self-reports do have the disadvantage that responses are vulnerable to tendencies like wanting to appear consistent or to provide socially desirable answers. The SRHI is no exception to that. Although the use of multiple items makes it less vulnerable to such threats compared to a single-item instrument, one should be careful when using it in the context of topics that are sensitive to such biases. In these cases, however, *all* self-report measures (including behavioral frequency estimates) suffer from the same problem, and one might prefer implicit or operative measures (Bassili, 1996; Greenwald & Banaji, 1995), such as the response-frequency measure (Verplanken et al., 1994). Such measures, however, cannot always be used because of practical considerations.

The measurement of habit strength has long been an underdeveloped issue. The use of behavioral frequency as a measure of habit, valid as it might

sometimes be, is clearly only a proxy for a true measure of habit strength. Importantly, behavioral frequency measures do not tap the heart of the habit construct: automaticity. The discussion of the role of habit in attitude–behavior relations has stalled, particularly because we did not have direct measures of habit strength. This might be read in the words of Eagly and Chaiken (1993), who wrote "the role of habit per se remains indeterminate . . . because of the difficulty of designing adequate measures of habit" (p. 181). The development of practical measures such as the SRHI might contribute to this discussion and stimulate the further development of theory and research on habit.

References

Aarts, H., & Dijksterhuis, A. (2000). Habits as knowledge structures: Automaticity in goal-directed behavior. *Journal of Personality and Social Psychology, 78*, 53-63.

Aarts, H., Verplanken, B., & van Knippenberg, A. (1997). Habit and information use in travel mode choices. *Acta Psychologica, 96*, 1-14.

Aarts, H., Verplanken, B., & van Knippenberg, A. (1998). Predicting behavior from actions in the past: Repeated decision making or a matter of habit? *Journal of Applied Social Psychology, 28*, 1355-1374.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*, 179-211.

Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology, 52*, 27-58.

Ajzen, I. (2002). Residual effects of past on later behavior: Habituation and reasoned action perspectives. *Personality and Social Psychology Review, 6*, 107-122.

Ajzen, I., & Fishbein, M. (1977). Attitude–behavior relation: A theoretical analysis and review of empirical research. *Psychological Bulletin, 84*, 888-918.

Ajzen, I., & Fishbein, M. (2000). Attitudes and the attitude–behavior relation: Reasoned and automatic processes. *European Review of Social Psychology, 11*, 1-33.

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, pp. 1-40). Hillsdale, NJ: Lawrence Erlbaum.

Bargh, J. A. (1996). Automaticity in social psychology. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 169-183). New York, NY: Guilford.

Bargh, J. A., & Gollwitzer, P. M. (1994). Environmental control of goal-directed action: Automatic and strategic contingencies between situation and

behavior. In W. D. Spaulding (Ed.), Integrative views of motivation, cognition, and emotion. *Nebraska Symposium on Motivation, Vol. 41* (pp. 71-124). Lincoln, NE: University of Nebraska Press..

Bassili, J. N. (1996). Meta-judgmental versus operative indexes of psychological attributes: The case of measures of attitude strength. *Journal of Personality and Social Psychology, 71*, 637-653.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.

Fredricks, A. J., & Dossett, D. L. (1983). Attitude–behavior relations: A comparison of the Fishbein–Ajzen and the Bentler–Speckart models. *Journal of Personality and Social Psychology, 45*, 501-512.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4-27.

Heckhausen, H., & Beckmann, J. (1990). Intention action and action slips. *Psychological Review, 97*, 36-48.

Hull, C. L. (1943). *Principles of behavior: An introduction to behavior theory.* New York, NY: Appleton-Century-Crofts.

James, W. (1890). *The principles of psychology* (Vol. 1). London, UK: Macmillan.

Kahle, L., & Beatty, S. E. (1987). The task situation and habit in the attitude–behavior relationship: A social adaptation view. *Journal of Social Behavior and Personality*, *2*, 219-232.

Landis, D., Triandis, H. C., & Adamopoulos, J. (1978). Habit and behavioral intentions as predictors of social behavior. *Journal of Social Psychology*, *106*, 227-237.

Mittal, B. (1988). Achieving higher seat belt usage: The role of habit in bridging the attitude–behavior gap. *Journal of Applied Social Psychology*, *18*, 993-1016.

Nisbett, R. E., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259.

Orbell, S., Hodgkins, S., & Sheeran, P. (1997). Implementation intentions and the theory of planned behavior. *Personality and Social Psychology Bulletin*, *23*, 945-954.

Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin*, *124*, 54-74.

Ronis, D. L., Yates, J. F., & Kirscht, J. P. (1989). 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, NJ: Lawrence Erlbaum.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Lawrence Erlbaum.

Schwarz, N. (1999). Self-reports: How the question shapes the answers. *American Psychologist*, *54*, 93-102.

Spector, P. E. (1992). *Summated rating scale construction: An introduction*. Newbury Park, CA: Sage.

Sutton, S. (1994). The past predicts the future: Interpreting behaviour–behaviour relationships in social psychological models of health behaviour. In D. R. Rutter & L. Quine (Eds.), *Social psychology and health: European perspectives* (pp. 71-88). Aldershot, UK: Avebury.

Tolman, E. C. (1932). *Purposive behavior in animals and men*. New York, NY: Appleton.

Towler, G., & Shepherd, R. (1991). Modification of Fishbein and Ajzen's theory of reasoned action to predict chip consumption. *Food Quality and Preference*, *3*, 37-45.

Trafimow, D., & Wyer, R. S., Jr. (1993). Cognitive representation of mundane social events. *Journal of Personality and Social Psychology*, *64*, 365-376.

Triandis, H. C. (1977). *Interpersonal behavior*. Monterey, CA: Brooks/Cole.

Triandis, H. C. (1980). Values, attitudes, and interpersonal behavior. In H. E. Howe, Jr. & M. M. Page (Eds.), *Nebraska Symposium on Motivation, Vol. 27* (pp. 195-259). Lincoln, NE: University of Nebraska Press.

Verplanken, B., & Aarts, H. (1999). Habit, attitude, and planned behaviour: Is habit an empty construct or an interesting case of automaticity? *European Review of Social Psychology*, *10*, 101-134.

Verplanken, B., Aarts, H., & van Knippenberg, A. (1997). Habit, information acquisition, and the process of making travel mode choices. *European Journal of Social Psychology*, *27*, 539-560.

Verplanken, B., Aarts, H., van Knippenberg, A., & Moonen, A. (1998). Habit versus planned behaviour: A field experiment. *British Journal of Social Psychology*, *37*, 111-128.

Verplanken, B., Aarts, H., van Knippenberg, A., & van Knippenberg, C. (1994). Attitude versus general habit. *Journal of Applied Social Psychology*, *24*, 285-300.

Verplanken, B., & Faes, S. (1999). Good intentions, bad habits, and effects of forming implementation intentions on healthy eating. *European Journal of Social Psychology*, *29*, 591-604.

Watson, J. B. (1914). *Behavior: An introduction to comparative behavior*. New York, NY: Holt.

Wittenbraker, J., Gibbs, B. L., & Kahle, L. R. (1983). Seat belt attitudes, habits, and behaviors: An adaptive amendment to the Fishbein model. *Journal of Applied Social Psychology*, *13*, 406-421.

Appendix A

The Self-Report Habit Index

Behavior X is something . . .

1. I do frequently.
2. I do automatically.
3. I do without having to consciously remember.
4. that makes me feel weird if I do not do it.
5. I do without thinking.
6. that would require effort not to do it.
7. that belongs to my (daily, weekly, monthly) routine.
8. I start doing before I realize I'm doing it.
9. I would find hard not to do.
10. I have no need to think about doing.
11. that's typically "me."
12. I have been doing for a long time.

Note. The items are accompanied by response scales anchored by *agree/disagree* and preferably should contain five or more response categories. In the present studies, 7-point (Studies 1 and 2) and 11-point response scales (Studies 3 and 4) were used. Some items may have to be reworded in line with the behavior under study.

Appendix B

Daily and Weekly Habits Included in the Second Session of Study 4

Daily habits	Weekly habits
Put out my shoes when I come home	Go out (3)
Make tea when I come home	Go into town on Friday afternoon
Make coffee	Travel by train (2)
Drink tea	Go shopping (2)
Drink coffee (2)	Inspect the refrigerator before shopping
Eat candies (2)	
Eat potato chips	Buy bread and milk
Eat an apple	Buy chocolate
Cook	Eat potato chips
Walk the dog	Cook dinner
Make the bed	Take a long shower
Brush my teeth (2)	Shave
Look into the mirror	Weigh myself
Wash my hands (2)	Take the newspaper from the mailbox
Read e-mail	Read the newspaper (2)
Open the mailbox	Exercise (2)
Look at my watch	Dance
Switch on music	Play basketball
Switch on the radio	Run
Switch on the radio when I wake up	Play guitar
Switch on the TV set	Go to church
Watch TV (2)	Call my friend
Watch a soap series	Call my mom
Watch GTST	Call home
Engage in TV channel surfing	Watch the news
Curse	Watch GTST
Smoke	Watch ONM
Go shopping	Clean up my room
Make a phone call	Vacuum
Go to classes	Clean the house
Take a seat on the left-hand side of the bus	Wash the dishes (2)
	Polish my shoes

Note. The numbers refer to the number of participants for whom this behavior was selected. Behaviors without numbers were presented only to 1 participant. GTST (*Goede Tijden Slechte Tijden* [Good Times, Bad Times]) and ONM (*Onderweg Naar Morgen* [Underway to Tomorrow]) refer to soap series that were popular on Dutch television at the time the data were collected.