Inducing a Stroop effect

Emmanuel M. Pothos & Katy Tapper

Department of Psychology

Swansea University

in press, Applied Cognitive Psychology

Running head: inducing Stroop

Word count: 4,976

Please address correspondence to E. Pothos or K. Tapper, Department of Psychology, Swansea SA2 8PP, UK. Electronic email may be sent at e.m.pothos@swansea.ac.uk or k.tapper@swansea.ac.uk.

Abstract

We examine the conditions that lead to Stroop interference for a meaningless linguistic label. Tiffany's (1990) model of drug abuse implies that individuals will respond more slowly to drug-related words compared to neutral words in an emotional Stroop task, because the former have many automatic associations (e.g., positive expectancies). To examine this proposal, we trained participants to associate a meaningless label with either one other word or several other words and examined the induced Stroop interference for these meaningless labels. In two experiments, and contrary to expectations from Tiffany's work, we observed greatest Stroop interference for the meaningless label with just one association. These results are discussed in terms of associative learning theory.

Introduction

The classic Stroop task is one of the most studied paradigms in psychology, with over 2,500 studies since the seminal work of J.R.Stoop (1935). There have been several proposals to explain the classic Stroop effect, such as the speed of processing of linguistic information relative to perceptual information (Schooler et al., 1997), the automaticity of processing linguistic information (Logan, 1980; MacLeod & Dunbar, 1988; cf. Meiran et al., 2002; Tzelgov, Porat, Henik, 1997), and connectionist modeling (Cohen, Dunbar, & McClelland, 1990). A central component in these accounts is a *conflict* between the information derived from the linguistic component of a stimulus and its perceptual component (for an overview, see MacLeod, 1991). Such a conflict *cannot* explain results in the emotional Stroop task, which is the focus of the present work.

In the emotional Stroop task, participants receive colored words, such that the meaning of the words has some emotional salience for the participants. For example, alcohol abusers will see words related to alcohol (as well as neutral words), anxious participants will see words related to their anxiety, etc. Interference in such versions of the Stroop task has been widely documented. For example, alcohol abusers will take longer to name the color of alcohol-related words, compared to neutral words (Bauer & Cox, 1998; Bruce & Jones, 2004; Cox, Blount, & Rozak, 2000), smokers will likewise be distracted by smoking-related words (Hogarth et al., 2003; Hogarth, Dickinson, & Duka, 2003; Mogg & Bradley, 2002; Waters & Feyerabend, 2000), and individuals restricting their food intake by food-related stimuli (Tapper et al., 2008). A key point is that a theoretical explanation for the (various version of the) emotional Stroop task has to be different from an explanation for the classic Stroop task (Algom, Chajut, & Lev, 2004; Cox, Fadardi, & Pothos, 2006; Williams, Mathews, & MacLeod, 1996). This is because in the emotional Stroop task there is no conflict between the perceptual properties of the words and their meaning, as in the classic Stroop. Rather, participants in emotional Stroop tasks get *distracted* by the meaning of the words, and have to delay processing the color of the words that is required to complete the task. The emotional Stroop effect has resisted a compelling theoretical explanation.

We briefly consider four explanations for the emotional Stroop, from Robinson and Berridge (1993), Waters and Green (2003), Cox and Klinger (1990), and Tiffany (1990). The last approach is purely cognitive, in that it postulates that the emotional Stroop effect arises from elementary associative learning processes between elements. Such elements do not have to be emotionally salient, so that the intriguing possibility arises that Stroop effects can be observed by associating neutral elements. Our purpose is to examine this theory in an experimental context with neutral stimuli, so avoiding complications from variables which are hard to measure (such as emotional salience). In other words, we are asking whether we can understand the emotional Stroop without reference to emotional salience. This exercise will also help relate associative learning theory with theories of substance abuse based on learned associations.

Incentive salience theory postulates that certain substances eventually become more desired, even if they are not necessarily liked (Robinson & Berridge, 1993). The brain becomes sensitized to information about the substance, so that such information becomes more attention ‘grabbing’. Accordingly, when alcohol abusers see an alcohol-related word in an alcohol version of the Stroop task, they will preferentially process the meaning of the word at the expense of the perceptual properties of the word. It is less clear, however, whether incentive salience theory can predict attention disengagement from word meaning, which would be needed to explain emotional Stroop results.

Waters and Green (2003) suggested that for alcohol abusers, the presence of alcohol-related stimuli elicits anxiety and so leads to a reduction in the efficiency of attentional processes. Hence, alcohol abusers are less able than light drinkers to ignore alcohol words in an alcohol Stroop task. Such an explanation cannot be a general one. For example, there are cases when attentional processes related to personally relevant stimuli are *more* efficient than for neutral stimuli. Using the dot probe task, Pothos et al. (2008) found that external eaters attended more rapidly to food-related cues compared to neutral ones (so that the relevant attentional process would be more efficient).

In the motivational approach to (e.g.) excessive drinking, Cox and Klinger (1990, 2004) suggested that our lives are organized around the pursuit of goals, each one of which corresponds to a current concern. A current concern will cause the person to direct attention towards and process more extensively cues in the environment relevant to the goal. Thus, an alcohol abuser performing the alcohol version of the Stroop task would be more distracted by alcohol-related stimuli, compared to neutral ones, leading to an attentional bias for such stimuli. An unexplained issue is the *persistence* of current concerns related to substance abuse: An alcohol abuser cannot just stop having a current concern for alcohol. This persistence suggests automatic processes (cf. Logan, 1988; Tzelgov, 1997), an observation which leads us to Tiffany's (1990) theory for substance abuse.

According to Tiffany (1990), the frequency of substance abuse results in the development of automatic associations between the substance and positive corresponding expectancies, elements of a person's routine that are related to the substance, related concerns etc. (Cox et al., 2006; McKenna & Sharma, 1995; Peretti, 1998; Stetter et al., 1995). Regarding the emotional Stroop, when (e.g.) an alcohol abuser encounters an alcohol-related stimulus, these automatic associations result in the activation of other information, leading to a temporary bottleneck in cognitive processing (cf. Collins & Loftus, 1975). Consistent with this view, Rather et al. (1992) reported that in a 'map' of the relations between alcohol-related concepts and other information for heavy and light drinkers, conceptual organization was more centered around the alcohol concepts for heavy drinkers compared to light drinkers (cf. Stacy, 1997).

Tiffany's (1990) proposal does not involve information about emotional salience, and so it can be examined with emotionally neutral stimuli. We trained participants to associate one meaningless word ('blib') with one real word and another meaningless word ('flet') with seven, related real words. The associations were trained over five consecutive days and learning was supported by explicit instruction (Hogarth & Duka, 2006). Subsequently, we examined the Stroop interference for the meaningless words. Which meaningless word led to the greatest Stroop interference, the one with a single association or the one with many associations? The latter result would support Tiffany's proposal, but the opposite result was observed.

We can consider Tiffany's theory and our proposed manipulation under the light of standard associative learning theory. According to one tradition, a stronger cue/target association should increase attention to the cue (Mackintosh, 1975; Rescorla & Wagner, 1972). Kruschke et al. (2005) implemented this idea in a connectionist model, in which the attentional salience of more reliable predictors was increased. Such a model accounted well for, e.g., the effects of blocking and highlighting. According to another tradition, more attention should be allocated to stimuli which have been involved in prediction errors (the idea being that a limited capacity agent has less to gain by attending to reliable predictors; Pearce and Hall, 1980). Consistent with this perspective, Wills, Croft, and Hodgson (2007) found a reduction in the amount of attention for reliable predictors using an EEG methodology and eye-tracking. Likewise, Hogarth et al. (2008a) found that unreliable predictors attracted more attention compared to predictors which reliably predicted either an outcome or its absence. With respect to the proposed manipulation, the association between a meaningless word and a single real word would be stronger so that, according to Mackintosh (1975) this meaningless word should lead to higher Stroop latencies. By contrast, according to Pearce and Hall (1980), in the case of the association between a meaningless word and several real words, the meaningless word would be a less reliable predictor of any particular real word; so, it is this meaningless word which should lead higher Stroop latencies.

The view that strength of association may direct attention in a Stroop task can be indirectly supported by the results of Tzelgov, Henik, and Leiser (1990). They examined Hebrew-Arabic bilinguals with a Stroop task, to find Stroop interference for both the native and the second language. The Stroop effect was larger in the native language, for which, presumably, associations between word forms and meanings were stronger (other research has also reported larger Stroop effects for participants' native language).

The emotional Stroop task has been the focus of extensive theorizing. In closing the introduction, we can consider why this has been the case. The emotional Stroop has had predictive value across a range of psychopathologies. Cox et al. (2002) showed that alcoholics in a treatment centre who showed an increased alcohol Stroop bias during their treatment were more likely to relapse three months later. Cox, Pothos, and Hosier (2007) found that alcohol Stroop bias predicted a reduction in the number of drinking days of excessive drinkers. Mogg et al. (1995) examined patients suffering from anxiety problems. Successful treatment led to the elimination of anxiety-related attentional biases. Understanding what aspect of cue reactivity the emotional Stroop measures will help explain these clinical findings and, possibly, lead to the development of cognitive-style interventions (Wiers et al., 2006). However, not all research supports the view of a causal link between substance abuse and corresponding attentional biases. For example, Hogarth et al. (2008a) reported drug seeking behavior even when a corresponding attentional bias had been abolished.

Experimental investigation

The experimental test of the above ideas involves two challenges. First, the training routine has to be extensive enough to warrant some semblance of Tiffany's (1990) postulated automatic associations. Note that the issue is not whether the associations are learned or not (they are very simple and they would be learned reasonably rapidly) but, rather, whether the associations are learned to a degree to which they can be considered automatic (and so lead to Stroop interference). The second challenge corresponds to how learning in the case of the single associate can be counterbalanced with learning in the case of multiple associates. Clearly, there are many ways to achieve such counterbalancing. Our approach in this work has been to consider two kinds of counterbalancing (examined in separate, between-participants conditions) and assess directly whether the form of counterbalancing makes a difference in our results or not (results show it does not).

Participants

Participants were 129 Swansea University undergraduate students, who took part for either course credit or a payment of £20. They were trained on five consecutive days (in nearly all cases Monday to Friday). The Stroop test was always on the last day. Sixty-nine participants were allocated to Condition 1 and 60 to Condition 2.

Materials and procedure

For both conditions, participants were taught to associate a meaningless label with one single (real) word or with seven related (real) words. In one version of the tasks, ‘blib’ was associated with ‘paper’ and ‘flet’ with ‘pliers, nails, spanner, rake, spade, hosepipe, and drill’ and in a second version ‘blib’ was associated with ‘paper, eraser, labels, desk, fax, Sellotape, and envelope’ and ‘flet’ with ‘pliers’. In both cases, the seven words were broadly related to each other, so as to capture the intuition that automatic links postulated by Tiffany (1990) presumably involve related concepts. (For example, for an alcohol abuser, there would be many automatic links between alcohol and concepts broadly related to alcohol and so to each other.) It is important to check that the ‘one’ and ‘many’ words associated with the meaningless labels are roughly equally emotionally salient, since any such difference could potentially lead to a difference in Stroop results. In pilot studies, we checked that the words in one category had roughly equal and neutral valence to the words in the other category. We do not describe these results in detail, since we also collected emotional valence data for each word from each participant individually; these results will be incorporated in the analyses.

To simplify the description of our research, henceforth *One label* refers to the meaningless label associated with one word, *Many label* refers to the meaningless label associated with many words, *One word* refers to the one word associated with a meaningless label, and *Many words* the many words associated with the other meaningless label.

The learning task involved five learning sessions on consecutive days. Each learning session involved a block of 96 trials, which were presented in a new, randomized order for each participant. Each learning trial in the task begun with a fixation cross for 350 ms, followed by a screen where a word or a meaningless label was presented on the left, an arrow was shown in the middle, and another word was presented on the right; for example, ‘Rake -------> Blib’. In the bottom right corner of the screen, the question ‘Go together or not?’ appeared, prompting participants to respond yes, if the association was an intended one, or no if it was not. The feedback was in the form of the words ‘Correct!’ or ‘Incorrect!’, shown in the middle of the screen for 700 ms., after which the next trial appeared. Whether a meaningless label appeared on the left or the right of the computer screen was counterbalanced (e.g.,, participants would both see a trial in which they would be asked about the association Blib ----> Paper and the trial Paper ----> Blib). Both participant responses and reaction times were recorded.

Condition 1 and Condition 2 differed in terms of how the counterbalancing for the One and Many associations was carried out. In Condition 1 there were the same number of trials with the One label as with the Many label, but, clearly there were many more trials with the One label in which the correct answer was ‘not goes with’ compared to the Many label. Condition 1 is explained in Table 1. In Condition 2 in each training session we showed only one occurrence of each possible negative trial (in which the answer was no), so that this condition predominantly involved positive trials (in which the correct answer was ‘yes’). Inevitably, in Condition 2 there were many more Many trials than One trials (Table 2). Thus, both conditions were designed to teach a certain association to participants, but emphasized different aspects of the learning problem.

--------------------------------TABLE 1, 2-----------------------------

After the first learning session, participants received a sheet of paper with the correct associations between the meaningless labels and the words. Note that Hogarth et al. (2008a; cf. Hogarth & Duka, 2006) reported that attentional biases for a target associated with a cue developed only for participants who were aware of the cue-target association. The instructions participants received for the learning task were as basic as possible. Participants were simply told that in each trial they would see a word on the left-hand side of the screen, and another on the right-hand side and that they would have to have to decide whether the two words 'go together' or not (by pressing the appropriate key). They were told that initially they would not know which words went with which, but that they would eventually learn through corrective feedback. No information was given about the Stroop task, which was going to follow the training sessions.

After the last training session on the fifth day, participants were told that they would see various words in different colors on the computer screen and that they must identify the color of the words as quickly and as accurately as possible. We used four colors, red, green, yellow, and blue; the four colors were employed approximately equally often. Participants first went through eight practice trials, in which they had to identify the print color of different color words (e.g., the word 'green' printed in blue etc.). Participants were told whether their response was correct or wrong after each practice trial. Following the practice trials, participants were asked whether they had any questions and, if not, they went through a Stroop task where the stimuli were the two meaningless labels and all the words that were associated with them. The stimuli were presented in a new, randomized order for each participant. No feedback was provided in these trials. Thus, the Stroop task consisted of 20 trials (two meaningless labels, plus the 8 real words associated with them, presented twice). For both the practice and the subsequent Stroop trials, each trial begun with a fixation cross which was replaced after 350 ms with the word that had to be responded to. The word was visible on the computer screen until a response was made. Reaction times were recorded with the response box provided by Cedrus corp. for Superlab, for better accuracy.

Finally, after the Stroop task, participants received a valence questionnaire, in which they had to rate various feelings for each of the words used in the associations. Specifically, participants were asked to rate on a Likert 1-7 scale, how much elation, happiness, pleasure, fear, sadness, and disgust they felt for each one of the words used in the learning task. In this way, we could control for any individual variation in the emotional valence of the words.

Results

All tests below have been run on reaction times for trials for which correct responses were provided and trials for which reaction times were not more than 6000 ms; responses which took more than 6000 ms were considered spurious. Regarding Condition 1, in the learning task, in Session 1, for some participants there were as many as 30% of trials for which reaction time was more than 6000 ms, but by Session 5 there were only a handful of trials which were eliminated. Regarding Condition 2, outliers were less than 5% for all participants for all sessions (training sessions or Stroop task sessions). Figure 1 shows the acquisition curves for the two conditions.

Predictably, the differences in the learning regime led to corresponding differences in the learning of the associations. Our objective is not to examine such differences in detail, apart from when they might impact on the corresponding Stroop task results. Note that we call a Stroop effect or Stroop interference the finding that ink-naming for one label takes longer than ink naming for the other label (by analogy with the fact that emotional Stroop interference refers to the longer ink-naming times for, e.g., alcohol-related words compared to neutral words).

--------------------------------------FIGURE 1--

Our experimental hypothesis relates to Stroop interference for the One or Many label, assuming that the associations with the real words have been highly learned. Accordingly, we added the error rate at training session 5 for both the One and the Many associations. This aggregate error rate ranged from 0% to 97.91%. We decided to eliminate all 13 (out of 129) participants with an aggregate error rate of 10% or more. Additionally, for another 14 participants there were irregularities in testing (e.g., noise from building work, mobile phone ringing, talking). These participants were also eliminated. All subsequent analyses were carried out on this reduced sample of 102 participants.

The main dependent variables were the Stroop interference for the One label and the Stroop interference for the Many label (that is, the color-naming response latency for the One/ Many label). We first examined whether there was a main effect of condition, by using between participants t-tests to compare reaction times for One and Many Stroop effects for Condition 1 and Condition 2; $t(100)<0.90$, $p>.37$ in all cases. This important result shows that the training regime did not have any influence on the induced Stroop effect for the meaningless labels. Also, we ran similar t-tests to examine whether the counterbalancing we carried out with respect to which words were associated with which meaningless labels had any effect (recall that for each condition we used two experimental scripts, reflecting different associations between meaningless labels and real words). All t-tests were also not significant ($t(100)<0.89$, $p>.37$ in all cases). We next examined whether there was any evidence of a speed-accuracy trade-off either in the learning results for Session 5 (the final session) or the Stroop tasks. We correlated RTs with error rates for the One and Many labels separately in Session 5 (respectively, $r=0.20$, $p=.046$ and $r=0.31$, $p=.002$) and for the Stroop results for the One label ($r=-.03$, $p=.79$) and the Many label ($r=-.11$, $p=.29$). There was no evidence for a speed accuracy trade-off (where correlations are significant, they are positive; a speed accuracy trade-off would be evidenced in negative correlations). The final preliminary check concerns Stroop interference for the real words associated with the One label (One word) versus Stroop interference for the real words associated with the Many label (Many words). Clearly, in this case we would expect no difference, since these were ordinary words chosen to be as neutral as possible. A paired-samples t-test confirmed this expectation ($t(101)=0.38$, $p=.70$).

Was there higher Stroop interference for the One label or the Many label? In standard emotional Stroop studies, a difference measure is computed for the RT of the words assumed to be more distracting versus the RT of the words assumed to be neutral. By analogy, we used a paired-samples *t*-test to compare the Stroop RT for the One and the Many label, which revealed the Stroop RT for the One label (M=754ms, SD=203ms) to be higher than the Stroop RT for the Many label (M=708ms, SD=151ms): $t(101)=2.36$, $p=.020$. Thus, it appears that the One label was more distracting than the Many label. Note that in Stroop experiments a control condition is typically compared to an experimental one. In our experiments, a control condition is implied by the fact that the two labels employed were meaningless words. Therefore, given the observed results, the Many label condition is effectively the control condition for the One label condition, so that, strictly speaking, Stroop interference applies only to the One label condition.

The above *t*-test illustrates the main finding, which readily generalizes in a more complete statistical test. We created a mixed-design ANCOVA with Stroop RT for the One and Many labels as a within participants factor, Condition 1 versus Condition 2 as a between participants factor, and two covariates. One covariate relates to whether some of our participants may have found the associates with the One label more emotionally salient compared to the associates with the Many label (despite the counterbalancing we did with respect to which words were associated with the One and the Many label, it is still possible that some differences might arise due to random individual preference). We computed an index of emotional valence for each word, as the sum of the positive feelings for a word minus the sum of the negative feelings for the word. The covariate variable was the difference in the emotional valence values for the One word minus the average of the emotional valence values for the Many words. A second covariate concerns the difference in speed of responding for the One associations compared to the Many associations in the last training session (Session 5). This covariate was computed as the difference in RT for trials with the One label minus trials with the Many label, in Session 5. The main effect for the difference between the color-naming latency of the One and Many labels was, as before, significant ($F(1,98)=7.99$, $p=.006$). The interaction with Condition was not significant ($F(1,98)=0.41$, $p=.53$). The covariates did not have any effect.

General Discussion

Tiffany’s (1990) model of substance abuse (and psychopathology more generally) provides an elegant and simple account of interference in the emotional Stroop. When, for example, an alcohol abuser sees an alcohol-related word a number of concepts linked to alcohol (such as ‘relaxation’, ‘the armchair where I usually drink’, ‘the fact that there is a bottle of gin in the cupboard’ etc.) are *automatically* activated, causing a temporary bottleneck in cognitive processing, which slows down reaction time. To examine this hypothesis under laboratory conditions, we trained participants to associate a meaningless label with one word and another meaningless label with many words. If this account of the Stroop were to be supported, we would have found a greater induced Stroop effect for the Many label compared to the One label. Our results were exactly the opposite, even taking into account potential differences in emotional valence of the words and learnability between the One and the Many associations.

The two conditions were meant to address different possible confounds in the training of the associations. In Condition 1 the One and the Many labels were presented to participants the same number of times, since it is possible that absolute frequency of (meaningless label) presentation is the most important possible consideration to control for (Table 1). However, a potential problem with this design is that in the learning task there were many more ‘goes with’ trials for the Many label compared to the One label; most of the learning trials for the One label were ‘does not go with’. It is possible that participants learned that the One label ‘goes with’ a certain word *and also* that the One label ‘does not go’ with certain other words (cf. Shanks & Darby, 1998), creating a potential imbalance in the complexity of the associations between the two labels. Accordingly, with Condition 2 we presented only one instance of each negative association for each of the meaningless labels (Table 2). In this way, the One label could no longer be said to involve more negative associations than the Many label and the only factor affecting the induced Stroop interference effect would be the number of positive associations: one vs. many. Of course, a problem with the design of Condition 2, is that there were many more learning trials involving the Many label compared to the One label. In sum, Condition 1 and Condition 2 had different (design) strengths and weaknesses. Crucially, there was no main effect of condition in the dependent variables of interest.

The finding that the induced Stroop effect for the One label was higher than for the Many label is consistent with the Mackintosh (1975) and Kruschke et al. (2005) proposals for associative learning, according to which a cue more strongly associated with a target will attract more attention. Likewise, Tzelgov et al.'s (1997) research suggests that when an association between words and meaning is stronger, then a corresponding Stroop effect would also be stronger.

An issue to consider here is whether we are justified in calling the observed results an induced 'Stroop' effect. In defining the emotional Stroop effect, Cox et al. (2006, p.444) pointed out that 'It reflects how performance suffers from selective attention to aspects of a stimulus that should be ignored in a task'. Researchers in the area have come to accept as 'Stroop effects' an increase in latency in a color-naming task, more or less regardless of how this increase arises. For example, Warren (1972) reported that a word on a Stroop trial should lead to higher Stroop interference if is primed from a previous trial. Klein (1964) also reported than any common word leads to some Stroop interference (noting, of course, that words whose meaning is related to color cause more interference; see also, Monsell, 2001). Likewise, in our experiments, we have a situation in which the color-naming latency for the One label increases relative to the color-naming latency for the Many label, presumably due to the different number of associations with each label (noting also that, because of how the results turned out, the Many label condition is implied to correspond to the Stroop control condition).

Overall, the induced Stroop paradigm provides some interesting possibilities for examining related theories of cognition (such as theories of associative learning) and with future work we hope to address such possibilities more carefully. Finally, note that other researchers have created artificial Stroop effects. For example, Schmidt and Besner (2008) have also explored an artificial Stroop effect, but in the context of the strength of associations between a word and a color; i.e., in the case of these investigators Stroop interference was the result of conflict. As far as we are aware, this is the first report of an induced Stroop effect analogous to the emotional Stroop effect.

The implications for Tiffany's theory are that, at best, this theory is incomplete, if we want to use the theory as an account of emotional Stroop (noting that Tiffany proposed his theory to explain substance abuse, rather than the emotional Stroop task). One challenge in Tiffany's theory is to explain how particular associations in a person's daily routine (such as ones involving alcohol and positive alcohol expectancies) become automatic and persistent to the point of disrupting a person's life, but other associations (e.g., feelings of pleasure with drinking water and water) less so. It is possible that associations about which we introspect become automatic more rapidly. As Hogarth and colleagues noted, unless a person is explicitly aware of an association, it is unlikely that such an association will affect attention (e.g., Hogarth et al., 2008a). Another possibility, motivated from Robinson and Berridge's (1993) theory, is that associations for which the cue and/or the target are more emotionally salient become automatic more rapidly (cf. Robinson &

Berridge, 1993). Finally, although there has been extensive consideration of the role of attention in associative learning, theories for the transition between a learned association and an automatic association have been less well developed. Understanding the development of automatic behavior in terms of theories of associative learning appears a promising direction for future research.

In closing, it appears that research on the induced Stroop effect has the potential of not only informing the current understanding of the emotional Stroop effect, but also potentially challenging current theories of automaticity and the circumstances that lead to automatic associations (cf. Logan, 1980, 1998; Kruschke et al., 2005; Tzelgov, 1997).

Acknowledgements

This research was partly supported from grant ReF05/1/010 from the Wales Office of Research and Development (WORD). We would like to thank Miles Cox, Dora Duka, Javad Fadardi, Bob Rafal, and two anonymous reviewers for their helpful suggestions. We also thank Rhian Worth for assistance with data collection.

References

Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional Stroop phenomenon: A generic slowdown, not a Stroop effect. *Journal of Experimental Psychology: General*, *133*, 323-338.

Bauer, D., & Cox, W. M. (1998). Alcohol-related words are distracting to both alcohol abusers and non-abusers in the Stroop colour-naming task. *Addiction, 93*(10), 1539-1542.

Bruce, G., & Jones, B. T. (2004). A pictorial Stroop paradigm reveals an alcohol attentional bias in heavier compared to lighter social drinkers. *Journal of Psychopharmacology*, *18*(4), 527-533.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97*(3), 332-361.

Collins, A. M. and Loftus, E. F. (1975). A spreading activation theory of semantic memory. *Psychological Review*, *82*, 407-428.

Cox, W. M., & Klinger, E. (1990). Incentive motivation, affective change, and alcohol use: A model. In W. M. Cox (Ed.), *Why people drink: Parameters of alcohol as a reinforcer* (pp. 291-314). New York: Gardner Press.

Cox, W. M., & Klinger, E. (2004). A motivational model of alcohol use: Determinants of use and change. In W. M. Cox & E. Klinger (Eds.), *Handbook of motivational counseling: Concepts, approaches, and assessment* (pp. 121-138). Chichester, United Kingdom: Wiley.

Cox, W. M., Blount, J. P., & Rozak, A. M. (2000). Alcohol abusers' and nonabusers' distraction by alcohol and concern-related stimuli. *American Journal of Drug and Alcohol Abuse, 26*(3), 489-495.

Cox, W. M., Pothos, E. M., & Hosier, S. G. (2007). Cognitive-Motivational Predictors of Excessive Drinkers' Success in Changing. *Psychopharmacology, 192*, 499-510.

Cox, W. M., Fadardi, J. S., & Pothos, E. M. (2006). The addiction-Stroop test: Theoretical considerations and procedural recommendations. *Psychological Bulletin*, *132*, 443-476.

Hogarth, L. & Duka, T. (2006). Human nicotine conditioning requires explicit contingency knowledge: is addictive behaviour cognitively mediated? *Psychopharmacology*, *184*, 553-566.

Hogarth, L. C., Mogg, K., Bradley, B. P., Duka, T., & Dickinson, A. (2003). Attentional orienting towards smoking-related stimuli. *Behavioural Pharmacology, 14*(2), 153-160.

Hogarth, L., Dickinson, A., & Duka, T. (2003). Discriminative stimuli that control instrumental tobacco-seeking by human smokers also command selective attention. *Psychopharmacology, 168*(4), 435-445.

Hogarth, L., Dickinson, A., Austin, A., Brown, C., Duka, T. (2008a). Attention and expectation in human predictive learning: The role of uncertainty. *Quarterly Journal of Experimental Psychology, 61*, 1658-1668

Hogarth L, Dickinson A, Janowski M, Nikitina A, Duka T (2008b) The role of attentional bias in mediating human drug seeking behaviour. *Psychopharmacology, 201,* 29–41

Klein, G. S. (1964). Semantic power measured through the interference of words with color naming. *American Journal of Psychology*, *77*, 576-588.

Kruschke, J. K., Kappenman, E. S., Hetrick, W. P. (2005) Eye gaze and individual differences consistent with learned attention in associative blocking and highlighting. *Journal of Experimental Psychology: Learning Memory and Cognition, 31*, 830-45

Logan, G. (1980). Attention and automaticity in Stroop and priming Tasks: theory and Data. *Cognitive Psychology*, *12*, 523-555.

Logan, G. D. (1998). What is learned during automatization? II. Obligatory encoding of spatial location. *Journal of Experimental Psychology: Human Perception and Performance*. *24*(6), 1720-1736.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163-203.

MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(1), 126-135.

Mackintosh, N. J. (1975) A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*, 276-298

McKenna, F. P., & Sharma, D. (1995). Intrusive cognitions: an investigation of the emotional Stroop task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1595-1607.

Meiran, N., Hommel, B, Bibi, U. & Lev, I. (2002). Consciousness and control in task switching. *Consciousness and Cognition*, *11*, 10-33.

Mogg, K., & Bradley, B. P. (2002). Selective processing of smoking-related cues in smokers: manipulation of deprivation level and comparison of three measures of processing bias. *Journal of Psychopharmacology*, *16*(4), 385-392.

Mogg, K., Bradley, B. P., Millar, N. & White, J. (1995). A follow-up study of cognitive bias in generalized anxiety disorder. *Behaviour Research and Therapy*, *33*, 927-935.

Monsell, S. (2001). Naming the color of a word: is it responses or task sets that compete? *Memory & Cognition*, *29*, 137-151.

Pearce, J. M. & Hall, G. (1980) A model for Pavlovian learning: variations in the effectiveness of conditioned but not unconditioned stimuli. *Psychological Review*, *87*, 532-552

Peretti, C. S. (1998). Anxiété et troubles cognitifs [Anxiety and cognition disorders]. *Encephale*, *24*, 256-259.

Pothos, E. M., Calitri, R., Tapper, K., Brunstrom, J., & Rogers, P. J. (2008). Comparing measures of cognitive bias relating to eating behavior. *Applied Cognitive Psychology*, *22*, 1-17.

Rather, B. C., Goldman, M. S., Roehrich, L., Brannick, M. (1992). Empirical modeling of an alcohol expectancy memory network using multidimensional scaling. *Journal of Abnormal Psychology*, *101*, 173-183.

Rescorla, R. A. & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.) *Classical Conditioning 2: Current Research and Theory*. Appleton-Century-Crofts, New York

Robinson, T. E. & Berridge, K. C. (1993). The neural basis of craving: An incentive-sensitization theory of addiction. *Brain Research Reviews, 18,* 247-291.

Schooler, C., Neumann, E., Caplan, L. J., & Roberts, B. R. (1997). A time course analysis of Stroop interference and facilitation: Comparing normal individuals and individuals with schizophrenia. *Journal of Experimental Psychology: General*, *126*(1), 19-36.

Schmidt, J. R. & Besner, D. (2008). The Stroop effect: why proportion congruent has nothing to do with congruency and everything to do with contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 514-523.

Shanks, D. R., & Darby, R. J. (1998). Feature- and rule-based generalization in human associative learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *24,* 405-415.

Stacy, A. W. (1997). Memory activation and expectancy as prospective predictors of alcohol and marijuana use. *Journal of Abnormal Psychology*, *106,* 61-73.

Stetter, F., Ackerman, C., Bizer, A., Straube, E. R., & Mann, K. (1995). Effects of disease-related cues in alcoholic inpatients: Results of a controlled "alcohol Stroop" study. *Alcoholism: Clinical and Experimental Research*, *19,* 593-599.

Stroop, J. R. (1935). Studies of interference in serial verbal reaction. *Journal of Experimental Psychology*, *18*, 643-662.

Tapper, K., Pothos, E. M., Fadardi, J. S., & Ziori, E. (2008). Restraint, disinhibition and food-related processing bias. *Appetite*, *51*, 335-338.

Tiffany, S. T. (1990). A cognitive model of drug urges and drug-use behavior: Role of automatic and nonautomatic processes. *Psychological Review*, *97*, 147-168.

Tzelgov, J. (1997). Specifying the relations between automaticity and conscious: A theoretical note. *Consciousness and Cognition*, *6,* 441-451.

Tzelgov, J. Porat, Z. Henik, A. (1997). Automaticity and consciousness: is perceiving the word necessary for reading it? *American Journal of Psychology*, *110*, 429-448.

Tzelgov, J., Henik, A., and Leiser, D. (1990). Controlling Stroop interference: evidence from a bilingual task. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *16*, 760-771.

Warren, R. E. (1972). Stimulus encoding and memory. *Journal of Experimental Psychology*, *94*, 90-100.

Waters, A. J. & Feyerabend, C. (2000). Determinants and effects of attentional bias in smokers. *Psychology of Addictive Behaviors, 14*, 111-120.

Waters, H., & Green, M. W. (2003). A demonstration of attentional bias, using a novel dual task paradigm, towards clinically salient material in recovering alcohol abuse patients? *Psychological Medicine*, *33,* 491-498.

Wiers, R. W., Cox, W. M., Field, M., Fadardi, J. S., Palfai, T. P., Schoenmakers, T., & Stacy, A. W. (2006). The search for new ways to change alcohol-related cognitions in heavy drinkers. *Alcoholism: Clinical and Experimental Research*, *30*, 320-331.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, *120*(1), 3-24.

Wills, A. J., Lavric, A, Croft, G. S., & Hodgson, T. L. (2007). Predictive learning, prediction errors, and attention: Evidence from event-related potentials and eye tracking. *Journal of Cognitive Neuroscience*, *19*, 843-854.

Tables.

Table 1. The structure of Condition 1.

Relative frequencies:	
One Label	
1 trial goes with	x 1
7 trials does not go with[1]	x 1
Many Label	
7 trials goes with	x 1
1 trial does not go with	x 1
Absolute frequencies[2]:	
One Label	
trial goes with	1 (word) x 2 (left, right)[3] x 3[4] = 6
trials does not go with	7 (words) x 2 (left, right)[3] x 3[4] = 42
Many Label	
trials goes with	7 x 2[3] x 3[4] = 42
trial does not go with	1 x 2[3] x 3[4] = 6

Notes. [1]'does not go with' refers to trials in which the correct answer was that the label and the real word do not go together. [2]Absolute frequencies refer to the total number of trials participants went through on each training day. [3] 'left. right' and the x2multiplier refer to the fact that each association was presented twice, one in which the meaningless label was shown on the left of the screen and another in which it was shown on the right. [4]Each learning session in this condition had three blocks of 32 trials, for a total of 96 trials.

Table 2. The structure of Condition 2.

Relative frequencies:	
One Label	
1 trial goes with	x 5
7 trials does not go with[1]	x 1
Many Label	
7 trials goes with	x 5
1 trial does not go with	x 1
Absolute frequencies[2]:	
One Label	
trial goes with	1 x 2[3] x 5 = 10
trials does not go with	7 x 2[3] x 1 = 14
Many Label	
trials goes with	7 x 2[3] x 5 = 70
trial does not go with	1 x 2[3] x 1 = 2

Notes. [1]'does not go with' refers to trials in which the correct answer was that the label and the real word do not go together. [2]Absolute frequencies refer to the total number of trials participants went through on each training day, which was 96 trials. [3]the x2multiplier refer to the fact that each association was presented twice, one in which the meaningless label was shown on the left of the screen and another in which it was shown on the right.

Figure

Figure 1. Acquisition curves for the associations with the One and Many labels for the two conditions. Error bars denote 1 SD; for clarity, we show only positive error bars for the One label curves and only negative error bars for the Many label curves.