Running head: GUIDE TO SEQUENTIAL PRIMING

A Practical Guide to Sequential Priming and Related Tasks

Dirk Wentura[1] & Juliane Degner[2]

[1]Saarland University, Germany

[2]University of Amsterdam, The Netherlands

Chapter prepared for B. Gawronski, & B. K. Payne (Eds.), ***Handbook of implicit social cognition: Measurement, theory, and applications*****. New York, NY: Guilford Press.**

Address correspondence to:

Dirk Wentura
Department of Psychology
Saarland University
D-66115 Saarbruecken, Germany
Tel.: +49-681-302-4642
Fax: +49-681-302-4049
E-Mail: wentura@mx.uni-saarland.de

Introduction

This chapter is about a family of simple tasks – sequential priming tasks – that have become increasingly popular in Social Psychology throughout the last decade (see Wittenbrink, 2007). The defining characteristic of these tasks is a common experimental procedure: Participants work through a sequence of trials in which two stimuli are subsequently or simultaneously presented but only one – the target – has to be processed with regard to a basic feature while the other – the prime – is task-irrelevant. Starting from this procedural commonality, the family splits into *two* dominant branches based on the kind of relation between prime and target and the kind of target processing task implemented. When it is the participants' task to classify letter strings as words or non-words (lexical decision task) and primes and targets vary according their semantic relatedness, we describe a variant of a *semantic priming paradigm*. When it is the participants' task to categorize targets as positive vs. negative (evaluation task) and primes and targets vary according their evaluative congruence, we describe a variant of a *response priming paradigm* (as we prefer to name it for reasons that will become clearer in the course of this chapter).

The basic result and dependent variable of interest is a difference in speed and/or accuracy between the conditions defined by the prime-target relationship within each trial. Typically, mean reaction latencies (error rates) in related or congruent trials are lower compared with those in unrelated or incongruent trials. Because primes are task-irrelevant and (*prima facie*) of no help for participants, such results reveal what is typically referred to as automatic processing of the prime (ignoring the complexities of the concept of "automatic" for a moment; see, e.g., Moors & De Houwer, 2006). Congruency effects in the aforementioned lexical decision task are assumed to shed light on the mental representation of these stimuli. This is *per se* an interesting phenomenon, but it becomes even more interesting if prime-target pairs of *a priori* known relatedness are replaced by pairs of assumed relatedness: e.g., primes denoting a social category and targets denoting presumed stereotypic traits (e.g., Wittenbrink, Judd, & Park, 1997). Congruency effects in the aforementioned evaluative categorization task are assumed to reflect automatic processing of the prime's valence. This is *per se* an interesting phenomenon but becomes even more interesting if primes of *a priori* valence are replaced by primes of assumed valence, e.g., pictures of black vs. white persons (e.g., Fazio, Jackson, Dunton, & Williams, 1995).

Although these two branches of tasks seem to be highly similar on first sight, we emphasize that they have to be accurately distinguished, as we elaborate on in this chapter. The purpose of this chapter is to describe different variations of sequential priming tasks and explain their different underlying mechanisms. We furthermore give an overview of typical designs, tasks, procedural features, and peculiarities of data analysis and potential application in social cognition research. We see this chapter as an introductory guide into sequential priming research that cannot answer all questions but gives valuable hints in how to do good priming research.

Mind the Gap! Semantic Priming versus Response Priming

The first branch of the priming family is known as the *semantic priming* paradigm, which was developed in the early 1970s (Meyer & Schvaneveldt, 1971). There are two typical variants, depending on the participants' task. The first one was already mentioned: Targets are words and (pronounceable) nonwords and the time to press a word (or nonword) key is recorded as the dependent variable (*lexical decision task*). In the second variant, targets (now words only) have to be read as quickly as possible (*naming or pronunciation task*), and the latency of naming is the dependent variable. The essential experimental manipulation in both tasks is whether the prime is semantically related or unrelated to the target. The typical difference between mean reaction latencies for related and unrelated pairs is explained by the activation of prime meaning, which facilitates encoding of related targets compared with unrelated targets (for exceptions to this rule, see later discussion; for reviews, see Lucas, 2000; McNamara, 2005; McNamara & Holbrook, 2003; Neely, 1991). The metaphor of spreading activation in a semantic network is typically used to illustrate the underlying mechanisms of such response facilitation effects (see, e.g., Collins & Loftus, 1975). Note, however, that this is not the only way to explain these effects (see, e.g., Masson, 1995; Ratcliff & McKoon, 1988).

The second branch of the priming family comprises different variants of *response priming*. The most prominent member of this branch in social cognition research is the *evaluative* (or *affective*) *priming* paradigm using the *evaluation task*, as developed by Fazio, Sanbonmatsu, Powell, and Kardes (1986). We have already outlined this paradigm: Participants categorize target stimuli with regard to their valence as positive or negative; the primes are positive or negative as well. Usually, primes are described as being evaluatively congruent or incongruent to the target, which is correct but somewhat misleading because it obscures the crucial difference to *semantic priming*. More appropriately, primes should be described as being congruent or incongruent *to the response* that has to be given to the target. The most straightforward explanation for a congruency effect in a *response priming* task is to assume that the irrelevant prime's response-relevant feature (e.g., the valence) is nonintentionally processed and prepares a response (see De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009; Klauer & Musch, 2003, for reviews): If the prime triggers the same response as the target, responding is facilitated (thus, responses are faster and errors less likely); if the prime triggers a different response, it conflicts with the response triggered by the target and this conflict has to be solved; thus slowing responses (or an error will more likely occur). At a more abstract level, response priming is structurally equivalent to the classic Stroop color interference paradigm (Stroop, 1935) or the Eriksen flanker task (Eriksen & Eriksen, 1974). For example, in the Stroop task, participants name the print color of words in congruent trials (e.g., the word *green* printed in green), incongruent trials (the word *green* printed in red), and neutral trials ("XXXX" printed in red). In incongruent trials, a response conflict occurs because the task-relevant word color suggests one response and the task-irrelevant word content suggests another. Before the response can be given, the response conflict has to be resolved (i.e., the task-relevant information has to be separated from the task-irrelevant information), resulting in slower and less accurate responses compared with congruent and neutral trials. Thus, in such response interference paradigms, a task-irrelevant feature (i.e., the word content in the Stroop task) or a task-irrelevant stimulus (i.e., the prime in the priming task) is automatically processed, triggering specific responses and thus interfering with responding in incongruent cases.

Beside the evaluative priming paradigm, there are other examples of *response priming* in social cognition research. The seminal article on automatic stereotyping by Banaji and Hardin (1996) is a good example for (1) a different version of *response priming* and (2) for the comparison of *response priming* and *semantic priming*, using the same materials and focusing on different facets of a research hypothesis (see also Wittenbrink, Judd, & Park, 2001). In their first experiment, the authors instructed participants to quickly categorize pronouns (e.g., *he*, *she*) with regard to gender. The pronouns were preceded by different categories of primes: Nouns that were related to gender either by definition (e.g., *father* vs. *mother*) or by normative base rates (e.g., *doctor* vs. *nurse*). Results indicated faster responding after generic gender-congruent prime/target pairs compared with gender-incongruent pairs. Given our prior definition, we have a clear example of *response priming*: The primes related to gender by definition can be categorized as male versus female on an *a priori* basis and can thus trigger the response "male" or "female." Thus, the observed priming effect can be interpreted analogous to evaluative priming: In the context of a gender classification task, a briefly presented gender-related word is automatically categorized as either male or female; this categorization helps or hinders the target categorization. This is *per se* an interesting phenomenon, but it becomes even more interesting looking at primes that were only indirectly gender connoted. Did *nurse* (*doctor*) *automatically* evoke a female (male) classification? Yes, it did: The congruency effect with normative base rate primes was significant as well.

In their second experiment, Banaji and Hardin (1996) changed the task from gender categorization to pronoun categorization (i.e., Is the target a pronoun or not?). Of course, trials with non-pronoun targets (e.g., *do, as*) were added. By this change in task instructions, we are faced by a version of *semantic priming*: The variation of interest (i.e., whether prime and target are matching with regard to gender or not) is completely orthogonal to the binary decision made with regard to the target. In this task, Banaji and Hardin found a relatedness effect for primes that were gender related by definition but not for primes that were only indirectly gender connoted. In two aspects, these priming effects differ from the effects found in the response priming version. First, automatic gender categorization of the generic gender related primes occurs even when no explicit gender classification is instructed, thus indicating a different degree of goal independence (or nonintentionality) of the process (see, e.g., Bargh, 1996). Second, the specific relation between prime and target counts: A concrete female-related word (i.e., *she*) is processed somewhat easier if it is preceded by a “female” noun. To add more emphasis to this point, imagine a further experiment in which the target set consists of adjectives (and adjective-like nonwords), and participants are instructed to categorize targets according to their lexical status (see, e.g., Wittenbrink et al., 2001). If adjectives are related to gender stereotypes, a priming effect would mean that encoding of a stereotypical “female” adjective (e.g., *caring*) is facilitated by a gender-related prime. Now the explanation of this facilitation has to draw on assumptions about the mental representation of stereotypes.

A transition between *response priming* and *semantic priming* as in the experiments by Banaji and Hardin (1996) can be made for evaluative priming as well. You can, of course, study evaluative congruency effects in a *semantic priming* design if you vary prime and target valence, for example, in a naming task (e.g., Bargh, Chaiken, Raymond, & Hymes, 1996; De Houwer & Hermans, 1994; Spruyt, Hermans, De Houwer, & Eelen, 2002) or a lexical decision task (e.g., Bessenoff & Sherman, 2000; Wentura, 2000). Note, however, that, in comparison to the *response priming* variant of evaluative priming, you add two essential assumptions. First, you assume that the valence of the prime is automatically processed even in a task context that does not ask for evaluation. Second and more importantly, you assume that the mental representation of valent stimuli is structured in a way that allows for encoding facilitation. This is an exciting research hypothesis on its own (see, e.g., Spruyt, Hermans, De Houwer, Vandromme, & Eelen, 2007; Wentura & Frings, 2008, for recent research).[1]

The devil inside: Strategic production of priming effects

Undoubtedly, the great interest in the priming paradigm is due to the belief that the priming effect could be a window to the inner structure of our cognitive (-affective) apparatus and the basic processes that operate within this structure. Why do we believe that? We treat participants in our studies as our confederates, as *persons* with the capability *to act intentionally* (a miraculous capability from a cognitive perspective), who are willed to tenaciously press keys in response to target stimuli in order to show a good performance in terms of accuracy and speed. As a mere side effect caused by the automatisms of their cognitive apparatus, and thus believed to be out of their control, their performance data reflect the priming effects we are interested in.

However, we should carefully rethink whether our data can alternatively be explained by strategic efforts of the participants. For the *semantic priming* branch, this question is closely associated with the seminal article by Neely (1977). What will participants do if they

---

[1] Note that the variation of prime and target valence in a binary categorization task (e.g., like the lexical decision task) allows for a further valence-specific process (see Wentura, 2000; see also Klauer & Stern, 1992).

realize the basic structure of a semantic priming experiment, that is, that a prime is often followed by a semantically related target? They might generate expectations about specific targets to appear after certain primes and, if an expected target appears, react faster compared with the unrelated condition. Neely pitted presumed automatic processes against expectancy-based priming by presenting in the majority of trials *a priori* unrelated pairs with a structure known to participants (i.e., they knew, for example, that whenever *body* was the prime, a part of a building [e.g., *door*] could be expected as the target; whenever *building* was the prime, a part of the body [e.g., *heart*] could be expected). Additionally, an unexpected minority of trials was composed of *a priori* related prime/target pairs (i.e., *body/heart*). In comparison to a neutral baseline, there was large facilitation for expected targets and large inhibition for unexpected targets. However, this pattern was moderated by the delay in onset of prime and target (stimulus-onset asynchrony [SOA]). It was strongest at an SOA of 700 msec and rather weak at an SOA of 400 msec. At an SOA of 250 msec, however, there was facilitation for *a priori* related targets (and no inhibition for unrelated but expected ones). What do we learn from this pattern? In *semantic priming* experiments, effects of *a priori* relatedness succeed even if participants are instructed to work against them, but only for short SOAs. Semantic priming effects found with long SOAs are more fragile and open for alternative explanations.

Another challenge of the interpretation of *semantic priming* effects in terms of encoding automatisms is posed by possible backward-checking strategies (e.g., Neely, Keefe, & Ross, 1989). Typically, the set of word trials (with half of them preceded by a related prime) is supplemented by the same number of nonword trials. Of course, nonword trials are preceded by word primes as well, however, typically without relationship to the nonword (e.g., *botter* as a nonword is usually not preceded by *bread*). Word as well as nonword targets initially activate entries in the mental lexicon (i.e., the word target *butter* as well as the nonword target *botter* activate *butter*). If a backward check yields the result that the activated entry has a relationship to the prime, it must be a word! We discuss how to prevent such strategies in the section Procedural Details.

The previous discussion applies exclusively to *semantic priming* designs. For *response priming* designs, we cannot claim to measure automatic effects just because we rely on a short SOA (e.g., 200 msec). Klauer, Rossnagel, and Musch (1997) conducted a study that can be regarded as an equivalent to the experiment by Neely (1977) for evaluative response priming. In this study, SOA as well as the proportion of congruent pairs was varied, which should lead to similar expectancies: If, for example, only 25% of all trials are congruent prime/target pairs, participants can expect a negative prime to be followed by a positive target (and vice versa) in the majority of trials. If evaluative priming follows the same laws as semantic priming, a manipulation of the relatedness proportion should determine the priming effects only at longer SOAs. That, however, was not what Klauer and colleagues found. Relatedness proportion moderated priming effects already at SOAs of 0 msec (i.e., prime and target were presented simultaneously one above the other) and 200 msec, with larger priming effects for larger relatedness proportions. However, even in the case of a minority of 25% congruent trials, the priming effect did not reverse, which might have been expected if participants intentionally prepare a response dependent on the prime category. Reversed effects (i.e., shorter reaction times [RTs] for incongruent trials) were only found at SOA = 1,200 msec.[2] What do we learn from these studies? First, given the analogy of evaluative priming to the Stroop paradigm (see earlier discussion), the result of relatedness proportion effects at SOA = 0 msec is not surprising; the pattern was known from the Stroop literature (e.g., Logan, 1980). Second, this moderation shows indeed that the process underlying response priming effects is

---

[2] We should note that at SOA=1,200 msec the effect no longer linearly depended on the relatedness proportion. Without going into detail here, this does not undermine our argument in the main text (see Klauer et al., 1997, for explanations).

somewhat more complicated than a simple automatism. Logan (1980) explains the moderation by relatedness proportion in the Stroop task by an attention-controlled adaptation of the weights of ink color and color word in generating a response. If the relatedness proportion is very low, the weight for the color word is reduced to a very low level and no longer interferes with ink color naming. If the relatedness proportion is high, a higher weight is given to the word content as an additional information source helping to achieve fast and correct responses. Third, this result is, however, not attributable to the same expectancy mechanism as the one tested by Neely (1977) because (1) an SOA = 0 msec does not allow for response preparation and (2) if participants would base their reactions fully on the contingencies, a reversed effects should have occurred for low relatedness proportions. The picture gets even more complicated because recent results suggest that evaluative priming can be influenced by participants' goals (Degner, 2009; Klauer & Teige-Mocigemba, 2007).

Interestingly, relatedness proportion effects were not observed for masked versions of response priming (Musch, 2000; Klauer & Musch, 2003), that is, for versions with primes for whom participants are at least subjectively unaware. This result indicates that supraliminal (unmasked) and masked versions differ somewhat in their processing characteristic.

Masked and unmasked priming

Masking procedures are used to prevent participants from becoming fully aware of the priming event and/or the prime content. For word stimuli, for example, a so-called sandwich masking procedure is often used: The prime is displayed very briefly (e.g., 10 - 40 msec) and embedded into letter strings serving as forward and backward masks. Typically, to the participant, the mask/prime/mask sequence appears as a brief flickering of letter strings, implying *subjective* unawareness of the prime. *Objective* unawareness is demonstrated by a so-called direct test: Participants work again through a sequence of the same kind of prime/target pairs (or a most comparable version of it) but are now instructed to categorize the prime instead of the target (see, e.g., Draine & Greenwald, 1998; see also, e.g., Dehaene et al., 1998, for a different version). The detection sensitivity parameter of signal detection theory (*d'*) is used as an index of (un)awareness. The most conservative evidence of subliminal priming consists of significant priming effects associated with a mean *d'* of zero (see, e.g., Schmidt & Vorberg, 2006). Others (e.g., Draine & Greenwald, 1998) suggest using regression analysis to provide evidence for priming in the absence of awareness: Regress the priming difference (incongruent RT minus congruent RT) on the *d'* of the direct test. If the intercept deviates significantly from zero, you can claim that there is priming at the zero point of the awareness measure. Of course, this only makes sense if the distribution of *d'* includes zero as a nonrare case (i.e., the scatter plot encloses the y-axis; see Klauer, Greenwald, & Draine, 1998, for a further problem).

For this kind of research in general and for applications in social cognition especially, a further differentiation is very important, although it seems to be a more technical detail at first sight. You can distinguish between experiments in which primes and targets are drawn from the same (small) set of stimuli (without, of course, presenting the same stimulus as prime and target in the same trial), such that the same stimulus is used as a prime in some trials and as a target in others, and experiments with primes that never appear as targets. For the former case, it is known that masked priming effects are robust (Draine & Greenwald, 1998). However, it seems as if they are caused by prime-induced retrieval of stimulus/response episodes that are created in trials where the prime was used as a target (see, e.g., Abrams & Greenwald, 2000). The latter case, however, is the more interesting one for social cognition researchers who want to use the masked affective priming paradigm as a tool in the attitude domain. Isn't it highly attractive to use primes that were never visibly presented (not to say evaluated) during the experiment? In fact, you can find effects in such experiments (see, e.g., Klauer, Eder, Greenwald, & Abrams, 2007; Wentura & Degner, in press, for basic research on this topic;

see, e.g., Degner & Wentura, 2009; Degner, Wentura, Gniewosz, & Noack, 2007; Frings & Wentura, 2003; Otten & Wentura, 1999; Wentura, Kulfanek, & Greve, 2005, for applications); however, they tend to be rather small.

The evidence for masked *semantic priming* effects is even more complex. There is evidence for *semantic priming* at very SOAs (e.g., the prime is overwritten by the target after 50 msec exposure; see Perea & Gotor, 1997; Perea & Rosa, 2002). For subliminal *semantic priming*, however, evidence is not as clear-cut as for *response priming*. Bodner and Masson (2003) found evidence for masked *semantic priming* (although direct prime identification was not perfectly at chance level), whereas Klinger, Burton, and Pitts (2000) contrasted response and semantic priming effects and found only evidence for the former but not the latter (while using the regression method of Draine & Greenwald, 1998). In the social cognition domain, for example, Wittenbrink and colleagues (2001) found stereotype-related priming by using masked presentation of category labels (i.e., black vs. white). The most important difference between the Klinger and colleagues and the Wittenbrink and colleagues experiments seems to be the massive repetition of the prime labels in the latter studies compared with the former. However, in our own experiments using repetition of category primes (albeit nonsocial ones) but a somewhat different masking procedure, we found *reversed* semantic priming (Wentura & Frings, 2005).

To summarize our report on masked versions of priming, we can state that *response priming* does produce masked effects. Even in the case with unpracticed primes (i.e., primes that are never openly presented throughout the course of the study), there is now enough evidence to rely on the existence of this (albeit rather small) effect. For *semantic priming*, one cannot take masked effects for granted. There is some evidence for priming effects, especially if one relies on the "weaker" criterion of subjective unawareness. Nevertheless, we strongly recommend pilot testing the procedure using standard materials.[3]

Design and Tasks

For both branches of the priming family, the design of a standard experiment is straightforward. We briefly discuss some details and variations with special consideration of social cognition research.

Basic Design of Semantic Priming

For semantic priming, semantically related and unrelated conditions have to be compared in a within-participants design (for the possible inclusion of a neutral priming condition, see later discussion). The gold standard of semantic priming experiments, however, is to present a given target only once to a given participant (to eliminate potential effects of target repetition). Thus, the design is basically a counterbalanced one (i.e., a Latin square): Two parallel lists of prime/target pairs are generated. Sample *A* receives pairs of list 1 as related pairings (related condition) and list 2 with a scrambled assignment of primes to targets (unrelated condition). For sample *B* the lists are presented the other way round. Thus, across the full sample, each target word appears equally often in the related and unrelated conditions. In social cognition research, you might only be interested in testing whether a certain set of stimuli *A* (e.g., first names associated with African Americans) prime a certain set of stimuli *B* (e.g., stereotypical trait words). It seems most straightforward to simply compare mean RTs to *B* targets preceded by *A* primes to mean RTs to *B* targets preceded by control primes (e.g., first names associated with Caucasian Americans) and to interpret a difference as a priming effect. This, however, is not legitimate because a difference might be due to an unspecific

---

[3] After finishing our chapter, a meta-analysis of masked priming was published (Van den Bussche, Van den Noortgate, & Reynvoet, 2009) that gives a more comprehensive review on that topic. Note that the terms "unconscious" and "subliminal" are used a bit less constrained by the authors than in the present paragraph.

main effect of prime set; that is, RTs to *any* target might be faster if preceded by words of set *A* in comparison to the control set (e.g., because words of set *A* are attention-grabbing; to illustrate, imagine an experiment using taboo words as primes). Thus, you are *obliged* to add a further set of target words as a control condition. Your priming effect corresponds to the interaction in a 2 (prime variation) × 2 (target variation)! Similarly, it is not recommended to use one list of related pairs and another list of unrelated pairs (with target words matched for some stimulus characteristics) for the whole sample. Any priming effect found might be due to subtle differences in word access or to subtle main effects of the two different prime lists. Thus, use the same list of targets preceded by related and unrelated primes.

In social cognition research, the material of interest (e.g., stereotypes) is often limited. Also, Latin square designs are not well suited for those social cognition studies that want to use priming indices as individual difference measures, because these make it preferable that all participants complete the same version of the task. Thus, targets have to be repeated in related and unrelated priming trials. We do not believe that repetition of targets (e.g., one time preceded by a related prime and one time by an unrelated one) is a problem *per se*. One should keep in mind, however, that results might become noisier. For example, repeated targets might reach a ceiling level of accessibility throughout the task or automatic target/response bindings might emerge such that primes cannot have any effect on target responses any more (see also McNamara, 2005, with regard to this topic). What kind of tasks can be used for *semantic priming* experiments (see also Table 1)?

Lexical Decision Task. As indicated early in the chapter, lexical decision is the classic and most frequently task for semantic priming research. Word and nonword trials are presented, and participants decide as quickly as possible whether the target is a legal word or not. Without challenging the status of this task, we want to mention three problems. First, nonword trials are in some sense "waste" trials because analyses are (almost always) restricted to word trials. Second, because typically nonword trials are semantically unrelated trials, participants are invited to use strategies that are not intended by the experimenter (see prior discussion on backward checking). Third, lexical decision processes are sometimes believed to be not elaborate enough for semantic analysis (see, e.g., De Houwer, Hermans, Rothermund, & Wentura, 2002).

Naming Task. The naming task has already been introduced in this chapter. The main advantage is that response-related explanations become rather implausible because a multitude of different responses have to be emitted by participants throughout the task. However, the task has some disadvantages in practical implementation: (1) a voicekey apparatus is required and (2) the responses must be monitored because one must constantly check whether the voicekey was triggered by the correct utterance (and not, e.g., by a wrong word or a cough). Thus, either an attentively listening experimenter has to categorize each response online as valid vs. invalid, or naming responses are recorded and arduously coded for accuracy afterwards.[4] A final point is to always be aware of possible confounds with regard to an unbalanced mixture of initial phonemes. If you were interested, for example, in simple differences to name positive versus negative words, you should carefully check whether one set of words contains more plosives (e.g., the “p” [p] in “positive”) while the other contains more fricatives (e.g., the “sh” [ʃ] in “shark”). Of course, this problem is not very severe in priming designs because throughout the experiment each target is presented equally often in related/congruent and unrelated/incongruent conditions.

---

[4] You might think of instructing participants to categorize each trial on their own with regard to whether the voicekey was triggered correctly. Of course, it should be clear that this is a less well-conducted experiment compared with the experimenter-controlled version. Whether a manuscript based on such experiments would find acceptance is your own risk.

Semantic Decision Task. De Houwer and colleagues (2002) tested for evaluative congruency effects in a *semantic priming* design by using targets that had to be categorized as either persons (e.g., *friend, enemy*) or objects (e.g., *gift, poison*). Advantages in comparison to a lexical decision are (1) presumably deeper processing of targets and (2) no waste trials (like the nonword trials in lexical decision). Banaji and Hardin's (1996) Experiment 2 mentioned previously provides a second example of semantic decision task. Of course, it has to be carefully considered whether a specific research hypothesis allows for such task variations.

Stroop Priming Task. An interesting but rarely used task is the Stroop Priming task (see Warren, 1972, 1974). It has its name as a result of a superficial similarity to the original Stroop color-naming paradigm: The target stimuli of a semantic priming task are presented in one of four or five colors and the color has to be named. It is assumed that mere variations in accessibility of concepts partly determine interference in color naming. Note that this cannot be inferred from the original Stroop studies because the original paradigm is dominantly a response interference paradigm (like response priming or the flanker task; see prior discussion). Nevertheless, Warren found increased color-naming latencies for targets that were preceded by semantically related primes. Rothermund and Wentura (1998) pointed out a potential advantage of this task, which can be made clear by a thought experiment. Imagine that you disambiguate the word *cancer* by a prime to refer either to the disease or to the astrological sign. Both priming conditions might result in increased accessibility of *cancer* in comparison to a neutral prime condition, thus principally enabling faster lexical decisions compared with the neutral prime condition. In the disease-prime condition, however, an additional effect might come into play: The possibly increased accessibility of a threat-related meaning of the prime might grab participants' attention and distract them from the task. The hypothesis for the response speed in a lexical decision task becomes uncertain: Increased accessibility of target meaning predicts fast RTs; attention grabbing of prime meaning predicts slow RTs. However, in a color-naming task, both hypothetical effects work in the same direction: They increase latencies. However, we hasten to remark that in a control condition, Rothermund and Wentura tried to replicate Warren's results by using associated prime/target pairs. This was successful when elaborate processing of the primes was assured (i.e., the prime word had to be memorized), which is rather unusual in semantic priming research. Without doing so, a more complex pattern of results was observed that made clear there are some intricacies associated with this task. In another application, Kawakami and colleagues (Kawakami, Dion, & Dovidio, 1999; Kawakami, Dovidio, Moll, Hermsen, & Russin, 2000) successfully used the task for stereotype assessment, but with a rather unusual long SOA of 1,000 msec (including a prime presentation of 950 msec). Thus, we can give no "easy-and-straightforward-to-use" recommendation for this task.

Further suggestions. In semantic priming research, the s*ingle-target procedure* has some reputation to reflect automatisms of memory retrieval: It consists of a sequence of binary categorizations (e.g., lexical decision) to single stimuli (i.e., there is no prime in the sense of a preceding task-irrelevant stimulus). The prime is simply the target of the foregoing trial (e.g., Shelton & Martin, 1992). Wentura and Brandtstädter (2003) used complete *sentences* as primes (see also Wentura & Greve, 2004, 2005). Finally, in the domain of text-processing psychology, *recognition priming* was successfully introduced: After reading some scenarios, participants went through a series of recognition trials with phrases out of the scenarios as targets. It was tested whether presenting one part of a scenario as a prime facilitates retrieval of another (see, e.g, Ratcliff & McKoon, 1978).

Basic Design of Response Priming

Basic *response priming design* comprises (at least; but see the *Procedural Details* section) the 2 × 2 variation of prime category and target category: The targets fall into two clearly defined categories, *A* and *B* (e.g., positive and negative). In basic research on response priming, the primes clearly differentiate into the same two categories *A* and *B,* and prime category and target category are orthogonally varied. In applying the design to research questions of social cognition (e.g., "implicit" prejudice), the primes differentiate into two sets as well, one whose exemplars are supposed to be automatically categorized as being a member of *A* and a control category that should not fall into the category *A*. In contrast to semantic priming, it is common to use rather small sets of stimuli. (That habit might have its origin in the structural equality to Stroop and flanker designs, which are always based on small sets of stimuli.) For *semantic priming designs*, it was easy to give an overview of different tasks that, despite their differences, test the same research hypothesis ("Do these primes facilitate encoding of those targets?"). For *response priming designs*, the task is intimately tied to the specific research hypothesis. For example, if you are interested in automatically evoked gender categorization of primes, use a male vs. female categorization task (see, e.g., Banaji & Hardin, 1996; see prior discussion). If you are interested in automatically evoked intergroup evaluations, use an evaluation task (see, e.g., Fazio et al., 1995). It is rather difficult to give an exhaustive list of tasks for the response priming paradigm. There are, nevertheless, three popular tasks in social cognition research that are variants of response priming designs. They are discussed in the following section (see also Table 1).

Evaluative (or affective) priming task. We have already mentioned the evaluative priming task several times, and there is no need to elaborate further on this task itself. Participants are simply instructed to categorize target stimuli as positive or negative as quickly as possible while ignoring the prime stimulus. A reasonable advancement is the use of response deadline or response window instructions.[5] If we consider response facilitation and/or interference as the driving force behind response priming effects, the accuracy of responding is in some sense the more “natural” dependent variable: If the prime supports the correct response, the likelihood of an error decreases; if the prime triggers the wrong response, the likelihood of an error increases. In somewhat more methodical terms, the probability of emitting a positive (vs. negative) response is a function of weighted target valence and weighted prime valence. Of course, if participants have enough time to process the target, they will finally press the correct response key in almost all trials, and facilitation or interference will dominantly show up in RTs. Thus, one trick to shift priming effects into accuracy data is to instruct participants to emit responses earlier than they would do in the nonrestricted case. In this regard, response window techniques can be very helpful (Draine & Greenwald, 1998). Response windows force participants to give extremely fast reactions within a predefined response window that is typically too short for high accuracy (e.g., within 300 – 450 msec after target onset). Thus, as participants are forced to perform with a more or less constant response speed, prime processing typically influences response accuracy, which becomes the main dependent variable. Simple response deadlines can have the same effect. Such techniques come with three further advantages: (1) The response window can be defined so close to the target onset that the target might not be fully processed when the response has to be given and thus prime influence on response accuracy is comparatively strong (i.e., the weight of the prime valence relatively increases; Draine & Greenwald, 1998; Wentura & Rothermund, 2003); (2) there is preliminary evidence that time pressure evoked by response deadlines impedes participants from implementing manipulating or faking strategies (at least in evaluative priming; Degner, 2009); and (3) besides any theoretical advantage of accuracy as the dependent variable,

---

5 Of course, the response window procedure is not exclusively bound to the evaluation task. Its use might be considered for any response priming task.

response windows shift potential priming effects into *one* dependent variable for all participants (i.e., an effect is not partitioned between accuracy, for fast participants, and response times – for accurate participants).

Affect Misattribution Procedure. Murphy and Zajonc (1993) developed a different version of affective priming. On each trial, a prime (i.e., an emotional face) was very briefly presented and directly followed (and thereby masked) by a *neutral* Chinese ideograph whose pleasantness had to be rated. They found that ratings were partly determined by the valence of the picture. Payne, Cheng, Govorun, and Stewart (2005) adapted this design to the assessment of automatic evaluation. In their *affect misattribution procedure (AMP)*, the prime is briefly (i.e. 75 msec) but visibly presented, followed by a blank screen (i.e., 125 msec) and an again briefly (i.e., 100 msec) but visibly presented Chinese ideograph as a target, which is finally overwritten by a mask. In a reduced binary categorization task, participants have to quickly decide whether the ideograph is less or more “visually pleasing than average” (p. 280). The priming effect in this paradigm consists of the actual evaluative responses to the neutral targets, that is the proportions of “more pleasant” versus “less pleasant” responses. Although participants are warned about the potentially biasing influence of the primes, the target categorizations typically reflect the valence of the primes: That is, “pleasant” responses are more/less frequent after positive/negative primes. This paradigm is in some sense the logical culmination of response priming: If we consider the probability of choosing one response category instead of the other as a function of weighted target valence and weighted prime valence (see prior discussion), the weight of target valence is almost reduced to zero if a neutral target has to be evaluated.[6] Further evidence for the role of response priming in the

---

[6] To prevent a misunderstanding, in line with Gawronski, Deutsch, LeBel, and Peters (2008), we do not assume the AMP to be driven by response *interferences* simply because a neutral target cannot trigger a clear positive or negative response. However, the affective priming effect in the evaluation task might have in common with the

AMP is the finding that the procedure also produces priming effects for nonevaluative variants of the task (e.g., Deutsch & Gawronski, 2009), which are difficult to explain with an affect misattribution mechanism. Results are impressive and the authors have done much to refute the criticism that participants follow a demand strategy (see Deutsch & Gawronski, 2009, for a comparison of AMP and standard affective priming).

Weapon Identification Task. With some reservations, we describe the weapon identification task (Correll, Park, Judd, & Wittenbrink, 2002; Payne, 2001) within the section on response priming. In a computer game, participants have to decide quickly whether (pictorially presented) target persons hold a weapon or a nonweapon object in their hands. In the former case, participants have "to shoot" (i.e., to press one key), and in the latter case they have "not to shoot" (i.e., to press another key). Target persons are of either black or white ethnicity. It is typically found that RTs and/or errors are decreased for black/weapon and white/nonweapon trials. On an abstract level, we are faced by a response priming design: A task-irrelevant feature (i.e., target ethnicity) varies orthogonal to the manipulation of a task-relevant feature (i.e., whether there is a weapon or not). What is different, of course, is that the mapping of the task-irrelevant feature to the categorization task is not as straightforward as, for example, in a basic affective priming experiment. (Only an extreme racist would cynically reconstruct the ethnicity variation in the experiment as a variation of "someone to shoot at" vs. "someone not to shoot at.") Thus, the paradigm invites one to think about ethnicity biasing early perceptual processes of the target object, such that a nonweapon object held by a black person is sometimes misperceived as a weapon. In recent studies, however, this was shown to be rather unlikely (Klauer & Voss, 2008; Payne, Shimizu, & Jacoby, 2005). According to

---

AMP that usually participants do not differentiate between the two sources of affect (i.e., prime and target). If they would try to do so, effects might be different from what is usually observed (see Wentura & Rothermund, 2003, for related arguments).

sophisticated analysis of data, response-bias explanations (and, therefore, the analogy to response priming) are more in line with the evidence.

Procedural Details

For the novice, programming a priming experiment asks for many procedural decisions for which no clear-cut rules can be found in the literature. For example, how many stimuli are needed? How often should each prime and target category be repeated during the task? How should timing routines be defined? Which control conditions are needed to interpret effects? How should instructions be formulated? We now provide some rules of thumb. We emphasize, however, that their application to specific research questions might depend on a variety of factors that we cannot discuss in detail in this chapter.

Target stimuli in semantic priming experiments

As mentioned, words are typically not presented repeatedly in basic semantic priming research. For applications in social cognition research, often this goal cannot be fulfilled because the set of materials for a given research hypothesis is typically quite small. However, any RT paradigm needs many trials (see later discussion). Thus, you have to find a balance between stimulus number (which ideally should be large), number of repetitions (which should be low), and overall number of trials (which should be large). Finally, finding the hypothesized, confound-free, and replicable semantic priming effect will justify your choices.

For a lexical decision task, nonwords should closely resemble words (e.g., generate nonwords from words by replacing one or two letters): If nonwords are too easily recognizable as nonwords (e.g., letter strings like *dfxsvg*), participants might use a quick “Is it a random string or not?” strategy and will thus not process the target stimuli beyond a perceptual level. In this case, semantic priming effects become highly unlikely. Finally, you might consider degrading targets (e.g., *bUtTer* or *%b%u%t%t%e%r%* instead of *butter*), because it is has been shown that semantic priming effects are enlarged if target processing is impeded (e.g., Becker & Killion, 1977; De Houwer, Hermans, & Spruyt, 2001; Williams, 1996).

To prevent backward-checking strategies (see prior discussion of strategic processes), you can think of using nonword trials that mimic the structure of the word trials (i.e., half of the nonword trials are "related" trials insofar as the prime and the word that was the basis of the nonword are related; see, e.g., Wentura & Frings, 2005).

Target stimuli in response priming experiments

In response priming experiments, it is important, of course, to choose target stimuli that are easily categorized with regard to the task at hand. For the evaluation task, use positive and negative words or pictures that can be unambiguously categorized as positive or negative according to some norm or pilot data. Note, however, that Chan, Ybarra, and Schwarz (2006) found *reversed* affective priming for high-frequency targets (while observing the usual positive priming effect for low-frequency targets). Beyond its value for understanding affective priming effects, these results should be taken as a hint that target characteristics might play a larger role than usually assumed.

On which dimensions should positive and negative target stimuli be matched? There is more than one answer to this question. The most conservative strategy is to prevent any confound. We would cling to that strict recommendation if we were writing about experiments exploring differences between positive and negative stimuli (e.g., whether one or the other are associated with longer RTs in the color-naming task), because any confound might be critical for interpreting differences in terms of valence (see, e.g., Delplanque, N'Diaye, Scherer, & Grandjean, 2007; Larsen, Mercer, & Balota, 2006, for possible confounds). However, in affective priming tasks, we test for the difference of RTs (and/or errors) to congruent and incongruent prime/target pairs, with the positive and negative stimuli entering both into the congruent and incongruent aggregates. Thus, confounds cannot distort the effect in a straightforward way.

At this point, we integrate an important issue that is often misunderstood: It is not legitimate to split the 2 (prime valence) × 2 (target valence) table into single effects, for example, to test whether affective priming occurs for positive but not for negative prime valence (or vice versa). Such single-difference analyses can be distorted by main effects of prime and/or target valence. To illustrate this, imagine a perfectly symmetrical crossover interaction pattern (that is the result of priming processes) in the 2 (target valence) × 2 (prime valence) table of means; for example, the congruent cells are associated with 580 msec and the incongruent ones with 600 msec. Now add main effects of prime and target valence to the table; that is, add 20 msec for all responses to targets following negative primes and 20 msec for all responses to negative targets. It *appears* as though priming works for one valence but not the other. Affective priming is the difference of mean RT to incongruent minus congruent priming.

If you stick to this principle, you need not worry too much about possible stimulus confounds. Nevertheless, if confounds are very blatant, they might be used by participants to facilitate responding (e.g., if positive targets are clearly shorter than negative ones), lowering the chances of finding reliable priming effects. It would be even worse if prime valence and target valence are both confounded in the same blatant way. In this case, the interpretation of congruency effects in terms of valence might no longer be justified because congruency vs. incongruency in the confounded feature might be responsible for the effect. Thus, try your best to prevent blatant confounds: For example, positive and negative words should be matched according to length and other perceptible characteristics. If you use the International Affective Picture System pictures (IAPS; Lang, Bradley, & Cuthbert, 2008), positive and negative pictures should not be distinguishable on the basis of perceptible characteristics, and the positive/negative distinction should not be replaceable by a more narrow pair of categories (e.g., *weapon* vs. *erotic*). Because arousal is the second dominant variable in the IAPS, it might be wise to match positive and negative picture according to arousal. Finally, if there are confounds that cannot be avoided, you might want to use regression analytical techniques (see later discussion) to demonstrate their (un-)importance.

Prime stimuli

For social cognition research, you will typically select several stimuli as primes that represent the concept or category in question. Asked to give a number, we would advise using four to ten stimuli for each prime category. However, there might be research questions that suggest a single stimulus as the cardinal one for the concept. For example, if you want to explore ingroup favoritism in the minimal group paradigm, the names of the ingroup and the outgoup might simply be the best (and only) available stimuli (Otten & Wentura, 1999). In such cases (see Frings & Wentura, 2003; Wentura et al., 2005, for other examples), you might want to mask primes in order to hide their massive repetition (see the research by Wittenbrink and colleagues, 1997; 2001, for a comparable example in the semantic priming domain).

Number of trials

Priming experiments are typically composed of a triple-digit number of trials.[7] The best planning strategy is to orient toward the smallest unit in terms of either (1) set size or (2) number of conditions of the design. An example of the former is an affective priming experiment using set sizes of 10 items (i.e., 10 positive and 10 negative targets as well as 10 positive and 10 negative primes). Accordingly, a block of trials would contain 40 trials with each prime and each target presented twice: once in a congruent and once in an incongruent pairing. In the second step, decide how many blocks are manageable in the context of your experiment (typically resulting in an aggregate of 100 – 300 trials). An example of the latter case is a study by Frings, Wentura, and Holtz (2007), whose design consisted of no fewer than 36 within-participants conditions. In such a case, align the trial number with the minimum number of trials needed for analyses. In our study, we did not want to fall below 10 trials per condition. Thus, the experiment eventually comprised 360 trials.

Would it have been acceptable to accept fewer than 10 trials per condition? Yes and no. Note that in the affective priming experiment outlined previously the number of congruent and incongruent trials would be between 60 and 140 (depending on whether the total number of trials would have been 120 or 280), which are typical values for RT-based experiments in cognitive psychology. Compared with this standard, 10 trials per condition is rather low. However, we have done semantic priming experiments (to confess, with some headaches) in which we have gone below this number (Wentura & Brandtstädter, 2003; Wentura & Greve, 2004). If you do so, be aware of the risk of finding data too noisy for significant priming effects and also the risk of triggering skepticism in reviewers.

For masked priming experiments, the number of trials per cell of your design should be even higher compared with supralimal studies. Masked priming might depend on a high number of prime repetitions before primes actually display the wanted activations. For example, in the aforementioned minimal group study (Otten & Wentura, 1999), reliable priming effects occurred only in the second of two blocks of 100 trials each. Moreover, these blocks were preceded by an additional practice block of 100 trials (as a result of using the response window technique; see prior and later discussions)!

Each priming experiment should include some practice trials (e.g., ten) in the instruction phase – their number depending on the complexity of the task – and some warm-up trials (e.g., 5 to 10) that precede the main trials unbeknownst to participants. Finally, filler trials can be interspersed if you need to hide the structure of the trials of interest.

---

[7] The AMP is an exception to the rule, because most studies have used 24 to 72 trials.

Neutral Priming Condition

In social cognition research, you might occasionally be interested in priming effects related to one prime category. For example, in evaluative priming experiments on intergroup bias, it might be of theoretical importance to separate spontaneous ingroup evaluations from outgroup evaluations. You might be tempted to compute separate ingroup and outgroup priming effects by subtracting responses to positive targets from those to negative targets and interpret positive scores as relative positive attitudes and negative scores as negative attitudes. As mentioned, such difference scores can be highly biased by various target effects (e.g., participants typically reacting slower to negative targets; participants reacting faster with the dominant hand). To compute priming effects, you always need some sort of baseline that helps you to separate main target effects from the priming effects of interest. For this reason, it might be wise to include a neutral prime condition into your experiment. Another reason for the inclusion of neutral primes is a cost/benefit analysis, that is, an analysis to separate facilitation (as a result of congruent/related priming) from inhibition (as a result of incongruent/unrelated priming). Although the use of neutral primes seems to be self-suggesting for this aim, things are a bit more complicated (see also Jonides & Mack, 1984). A simple consideration shows that the comparison of mean RTs for congruent and incongruent trials with the neutral baseline is ambiguous: Assume that the mean RT for congruent trials is 500 msec and the mean RT for incongruent trials amounts to 530 msec. If the mean RT for the neutral baseline is 530 msec, it seems as if priming was entirely due to facilitation mechanisms. Now assume that prime valence is associated with a main effect: For examples, besides facilitating negative responses for most of the trials negative primes capture attention in some trials, thereby increasing RTs for positive as well as negative targets. This main effect will increase mean RTs for the congruent as well as for the incongruent conditions. To extend our example, if this increase amounts to 30 msec, we would be faced by a result suggesting that priming is entirely due to inhibition (because the mean RTs for congruent trials and incongruent trials are now 530 msec and 560 msec, respectively).

What would be the most appropriate baseline condition? You might decide to use words or pictures that have been rated as neutral regarding the dimension of interest. In this case, be aware that neutral as well as ambivalent stimuli reach neutral average ratings in a pretest. Thus, if you don’t want to run the risk that your neutral primes are actually positive for some participants and negative for others (thereby highly biasing your baseline), take the variance of ratings or rating differences between groups of participants into account when choosing your stimuli. Alternatively, you may decide to use some nonsense primes, for example, nonwords, random (consonant) letter strings, scrambled pictures, or nonsense patterns for a neutral baseline. Of course, the problem of nonneutrality is not completely ruled out (e.g., nonsense might act like negative stimuli because of contrast effects).

Adding standard priming materials

Priming effects are well-documented phenomena. However, it would be a misconception to think that priming effects are as easy to replicate as, for example, the Stroop interference effect. Although *response priming* effects with the evaluation task are rather robust, we occasionally find zero effects. Also there are reports in the literature on reversed effects (Chan et al., 2006; Glaser & Banaji, 1999; Wentura & Rothermund, 2003). The world of *semantic priming* effects is even more complex: Using clearly associated materials (e.g., *bread/butter*), you will certainly find a reliable semantic priming effect. However, if you enter the literature on nonassociative semantic priming (e.g., whether category coordinates prime each other), finding robust effects cannot be taken for granted (for reviews, see Hutchison, 2003; Lucas, 2000). Thus, if you plan to use priming experiments to test specific hypotheses (e.g., “Is there evidence for automatic ageism?”), especially if you do so for the first time, it is important to verify that the specific procedure and instructions you use actually do allow for the occurrence of the priming effects in question. For example, in using the evaluation task for assessing automatically activated valence, we typically use additional control primes of clear positive and negative valence (according to norm data). In semantic priming experiments, prime/target pairs whose semantic relatedness has been demonstrated in previous experiments might be included.

Timing Routines

Ideally we would provide standards on presentation durations and timing intervals. Unfortunately, there is little research systematically exploring these procedural details, except for SOA. Therefore, any value given here is not carved in stone but simply reflects what we see as typical. We do not rule out that a systematic exploration of some of these values might lead to discoveries of theoretical interest.

A trial typically starts with a fixation stimulus (e.g., a”+” presented for 500 msec). Its only function is to orient the participant's attention toward the location of prime and target presentation and inform about the beginning of a new trial. Subsequently, the prime is presented typically with a duration varying between 100 and 300 msec (of course, we focus on supraliminal priming at the moment). Primes are replaced by either a blank screen or the target directly. The most important, and best explored, timing parameter in priming research is the SOA between prime and target. We have already reported findings related to SOA and strategy effects. At this point, we want to give some SOA recommendations for finding robust priming effects. For *response priming*, an often-used SOA is 300 msec (with a prime presentation of 200 msec), which is the value introduced in the original experiments by Fazio and colleagues (1986). Using this SOA, you will typically obtain effects. A recommendation, however, is to use SOAs less than 300 msec (e.g,, 150-200 msec; see also Klauer & Musch, 2003): Klauer and colleagues (1997) as well as Hermans, De Houwer, and Eelen (2001) tested for several SOAs and found reliable affective priming at SOA = 150 msec but not at SOA = 300 msec (or above).

In *semantic priming* studies, it is common to use SOAs of 250 msec or less because of the automaticity issue (see prior discussion). A typical, and unfortunately misled, assumption is that short SOAs less than 300 msec guarantee automatic processing (e.g., that participants do not strategically manipulate priming effects). However, if participants take their time to see and profoundly process the target and give a response (e.g., after 2000 msec), you cannot assume that the response would be free of strategic influences. A further means to reduce strategic influences might be the implementation of a response deadline or window (see, e.g., McKoon & Ratcliff, 1986; see also Degner, 2009). Thus, whereas targets are typically presented until the participant presses a response key, response deadline and response window techniques shorten target presentation (e.g., to 500 msec) and oblige participants to respond within a restricted time frame. A final word on timing routines: Response window timing is typically *adaptive*. That is, the center of the window adapts to the current performance of a participant to ensure a rather large error rate. Procedural details are reported in Draine and Greenwald (1998).

Studies vary in whether direct accuracy feedback is given or not. In response window experiments or any experiment with speed instructions, you typically do not give instantaneous error feedback. For an accuracy-stressing experiment, that is, when your main dependent variables are reactions latencies, you might consider direct error feedback and forced corrections (i.e., the next trial starts only after the correct response has been given). This not only reminds participants of accuracy instructions but has the additional advantage that participants do not start pressing the response keys randomly hoping the experiment would be over soon. As an intertrial interval, values between 1,000 and 2,000 msec are typically implemented.

Masking

To prevent participants from becoming (subjectively or objectively) aware of the prime event, use brief prime presentations and masks. Duration of prime presentation is usually between one and three refresh cycles of the screen (see later discussion), that is, for example, between 14 and 43 msec (12-35 msec) for a 70-Hz (85-Hz) screen. Typically, sandwich masking procedures are implemented, in which the briefly presented prime is embedded into a forward and backward mask. Word stimuli are typically easy to mask; a random sequence of letters (consonants) will do. The advantage is that participants often get the impression that the computer would shuffle through the letters of the alphabet before presenting the target stimulus and are thus less suspicious about the priming event. It is more difficult to choose good masking stimuli for pictures, especially for faces. Be aware that there is a large amount of basic research on subliminal perception that you should rely on for finding the best masking procedure (e.g., Breitmeyer & Ögmen, 2006). For example, to mask facial expressions, you might consider using spatially quantized face pictures (e.g., Bachmann, Luiga, & Pöder, 2005). Note that masking success is a function of the specific prime and mask stimuli as well as presentation duration. You thus should carefully pilot test whether the masking procedure you choose fulfills your expectations regarding subjective or objective unawareness of the priming event (see prior discussion).

Screen synchronizing

The shorter the durations, the more a technical aspect comes to the fore, especially if you use masked priming. To present a stimulus on a screen means (in technical terms) that the stimulus is stored into a screen buffer. With a certain frequency (typically between 70-100 Hz), the cathode ray of the traditional type of computer screen starts in the upperleft corner of the screen to display the content of the buffer on the screen. If you do not synchronize the cathode ray with your “present” command, the following scenario will unfold. Assume that you have a "present-the-prime" command in your experimental software, followed by a "delay" command of 20 msec, and finally a "present-the-mask" command. Additionally assume that your screen is updated 70 times per second, that is, one cycle of the cathode ray lasts approximately 14 msec. Imagine the following: First, your "present-the-prime" command alters the screen buffer just before the cathode ray arrives at the location of the prime. Thus, the prime will be almost immediately presented on the screen. The "present-the-mask" command will alter the buffer after 20 msec, that is, when the cathode ray has written the prime a second time. Thus, the prime will be presented for 28 msec. Second, your "present-the-prime" command alters the screen buffer just after the cathode ray had passed the location of the prime. Thus, the prime will appear on the screen after approximately 14 msec has elapsed. After a further 6 msec the buffer will be changed according to the "present-the-mask" command, such that in the next cycle of the screen the mask is presented. That means that your prime will be on the screen for a total of only 14 msec in total. To prevent these unpredictable variations, you have to synchronize updating of the buffer with the cathode ray refresh signal (i.e., that synchronous to the "present-the-prime" command, a refresh cycle will start). Typically, software packages for creating experimental programs include this feature. Furthermore, choose presentation times that represent a multiple of the cycle implied by the refresh rate (e.g., presentation times 14 msec, 28 msec, etc., for a cycle of 14 msec implied by a refresh rate of 70 Hz). It is recommended to use a "delay" after the "present-the-prime" command that is some milliseconds less that value (e.g., 10 msec instead of 14 msec, 24 msec instead of 28 msec) to be sure that in the follow-up cycle the mask will be presented. Some programming software packages for creating experimental programs already include corresponding features (e.g., by defining stimulus presentation times in terms of frames instead of milliseconds).

Randomization

According to standards of experimental cognitive psychology, priming experiments should consist of fully randomized trial sequences as well as full counterbalancing of all experimental conditions. Randomization is needed if the main result of your study is a general one, that is, one that should hold for the average participant because it precludes alternative explanations of results (e.g., based on specific trial sequences). However, it might be necessary to constrain randomization. For example, direct repetitions of stimuli should be avoided (it would lead to artificially large facilitation effects) and each prime category, target category, or prime/target condition should be presented equally often, without too many direct repetitions. If the most important result of your study is concerned with individual differences, it might be wise to use the same randomized sequence of trials for all participants to reduce error variance that is due to different randomized sequences. (If you conduct an intelligence test, you typically do not shuffle the items for each participant either.) You might even renounce counterbalancing. Of course, without randomization and counterbalancing, main effects cannot be straightforwardly interpreted anymore. It is as Banse (2001, p. 517) has phrased it: “One has to decide whether to optimise the design … for the analysis of treatment effects or for the analysis of individual differences, one cannot have it both ways.”

Instructions

In typical priming experiments, participants are simply informed about the presentation of two stimuli, and they are asked to perform a given task on the second one. Thus, the instruction states (either implicitly or explicitly) to ignore the prime. Apart from this standard, there are some studies that explicitly instruct participants to process the prime. We have mentioned the Stroop priming task in which Rothermund and Wentura (1998) found straightforward results only if the prime had to be memorized. It is interesting to note that Fazio and colleagues (1986) followed the same strategy introducing the evaluative priming paradigm. In their later studies using evaluative priming for assessing social attitudes (Fazio et al., 1995), they even used a prime-related task to disguise their intention: Prime faces had to be processed for a later memory task. In fact, there is evidence from different subdomains of priming research that processing of the prime matters. In "implicit" attitudes research, Olson and Fazio (2003) found different results depending on whether primes had to be categorized or not. In basic research on semantic priming, there are results that show that semantic priming disappears if the prime is processed very superficially (see, e.g., Stolz & Besner, 1999; but see Heil, Rolke, & Pecchinenda, 2004). Thus, there is a need to explore these issues in more detail.

In the instructions you should emphasize speed of responding while simultaneously conserving a rather high rate of accuracy (unless you use the response window technique; see prior discussion). Typically, you will achieve the appropriate speed/accuracy trade-off by mentioning to participants that they will inevitably make some errors if they are as fast as they should be.

Details of analyzing priming data

Preparation of data

RT data are characterized by some peculiarities that ask for careful data preparation. This includes trimming of individual data and sample screening for outlying participants. The first is important because not all priming trials result in valid RTs. The latter is important because any priming paradigm requires that participants intentionally fulfill the task at hand (and delivers the priming result as a by-product). If there are indications that participants did not fulfill instructions, it might be wise to discard their data from analyses.

It is a *must* to screen distribution of RT data for outliers. On the left of the RT distribution (the "fast side"), you often have no or few outlying values according to distribution-based criteria. However, an RT of, for example, 90 msec cannot be valid in a lexical decision task. Thus, you have to discard extremely short RTs. Typically, a value between 150 and 300 msec is chosen as a minimum inclusion criterion. On the right of the distribution, there are typically some outlying values that reflect lapses of attention or a second thought about the decision. In different articles about RT studies, slightly different solutions to this problem are offered. Often, an experimentwise criterion is defined, marking implausible long RTs for the task at hand. For example, RTs greater than 1,500 msec are certainly invalid in a task that is associated with mean RTs in the range of 500 to 700 msec. A good solution to find this criterion for a given task is to accumulate all valid RTs (i.e., RTs associated with correct responses) of all participant, and to discard either all values that are “far out” or “outside” (i.e., values that are 3 or 1.5 interquartile ranges, respectively, above the third quartile of the distribution Tukey, 1977).[8] The advantage of this procedure compared with one that defines outliers with regard to standard deviation units above the mean is obvious: The extremeness of outlying values themselves do not distort the definition procedure. In case of experiments with different conditions that are *a priori* associated with huge differences in mean RT (e.g., words and nonwords in a lexical decision task), it might be wise to define outlier criterions on the basis of the conditionwise RT distributions. Of course, each participant performs at his or her own speed. Thus, it is consequent to define outliers for the individual distribution as well, for example, based on the criteria proposed by Tukey (see prior discussion). Finally, you might combine sample-based and individually defined criteria.

What is a typical rate of outliers? Instead of giving a rule of thumb, we have simply checked our last 10 publications on priming experiments (including 23 experiments). Outlier rates were between 0.29 and 5.80%. Note, however, that Ratcliff (1993) concluded from simulation studies that it might be wise to discard up to 10% of the upper tail of the RT distribution. (Because we do not see such high rates published often, it might be wise to cite Ratcliff to convince skeptical reviewers.)

---

[8] “Far out” (“outside”) are the values marked by a * (a °) in a SPSS box-plot.

Calculation of mean RTs is based on all nonoutlying RTs that are associated with a correct answer. To adjust for the skewed distribution of RTs, you can transform raw data (for suggestions, see, e.g., Fazio, 1990; Ratcliff, 1993), with the log-transformation being most frequently used. Note that instead of using means of transformed, trimmed RTs, you can alternatively use the individual medians as the repeated measures (see Ratcliff, 1993, for recommendations; see Miller, 1988, for a warning given some specified conditions).

Data of participants with extreme error rates and/or extreme mean RTs should be discarded (see prior discussion). Exclusion criteria are, of course, dependent on the concrete materials and task used in a specific experiment. You are well advised to box-plot participants' mean reaction latencies and mean error rates and exclude participants with outlying or extreme values.

Analysis of Data

We briefly consider some aspects of data analysis that might be useful in analyzing priming data (and other RT data).

Difference Scores. Personally, we rarely use mean RTs or error rates, preferring instead to use priming difference scores as main dependent variables of a priming experiment (i.e., mean RTs or errors for unrelated/incongruent trials minus mean RTs or errors for related/congruent trials). Especially if the priming factor is embedded into a more complex design, this facilitates understanding of results. Although it should be clear to anyone who handles such data, we state it explicitly: Analyzing such difference scores is formally equivalent to analyzing the nonreduced plan (with, e.g., the F test for the effect of prime/target relatedness in the nonreduced analysis being equivalent to the one-sample t test of whether the mean difference score is significantly different from zero). What is the advantage? First, you eliminate the reporting of main effects and interactions that do not involve your priming factor. Second, thinking about differences and their significant deviance from zero reintroduces the idea of directional tests. Third, it naturally links experimental and correlational designs because the difference scores can be correlated with individual differences variables.

Individual regressions. In the world of between-participants designs, it is wellknown that nominal-scaled predictors (e.g., coding variables for experimental designs) and interval-scaled variables as well as their interaction terms can be introduced into a single regression model. Less well known is that this applies for repeated measures designs as well (like priming designs; see Lorch & Myers, 1990). The most simple and straightforward application is to (1) calculate regression models for each individual data set (with trials as the unit of analysis) and (2) test whether the means (across participants) of the regression weights are significantly above or below zero. For example, if you want to test whether RT in an evaluation task is better predicted by valence congruency or by a (continuous) index of semantic relatedness of prime/target pairs, you can easily do so by testing a model with valence congruency and semantic relatedness as predictors (see Wentura & Degner, in press). If the mean of the regression weights for affective congruency is significantly above zero but the mean for semantic relatedness is not, you can infer that affective congruency is the better predictor of reaction latencies. Lorch and Myers (1990) describe how to calculate these tests in a single step by using the complete N (number of participants) × M (number of trials) data set. More advanced and up to date would be the use of *hierarchical linear models* (HLM; see later discussion). The basic rationale, however, remains the same.

Reliability. As mentioned repeatedly, priming indices are often used as individual differences measures. This raises the issue of reliability of priming indices. If you follow classic test theory, you calculate two parallel priming scores, correlate them, and adjust according to the Spearman-Brown formula. Two parallel scores might be calculated by splitting the entire trial sequence into two halves. However, be aware that the first half and the second half of a priming experiment might be different with regard to the processes involved (see earlier discussion of *number of trials*). Just as well, you should be cautious to use an odd/even splitting. If there are autocorrelations (i.e., the response in trial *n*+1 is not independent from trial *n*), the split-half correlation is biased. An alternative might be to use several random splittings of the trial sequence and report the average reliability score that results. As an alternative to the Spearman-Brown adjusted split-half correlation, you can calculate more than two independent scores, which you then consider as item scores for calculation of Cronbach's alpha. For example, in an affective priming study with 10 positive and 10 negative targets that are repeatedly presented, you assign each negative target to one positive target and calculate priming scores for each positive-negative pair (see, e.g., Payne, Cheng et al., 2005, for a comparable solution). Reliability indices are typically rather low for indirect measures (but see Payne, Cheng et al., 2005, for the AMP). This seems to be especially true for priming measures (see, e.g., Banse, 2001; Cunningham, Preacher, & Banaji, 2001; Olson & Fazio, 2003; Wentura et al., 2005, who report reliability indices). A simple consideration might help to identify some aspects of the problem. By using attitude object-related primes (e.g., pictures of blacks vs. whites) in an affective priming task, the variance in the priming effect is interpreted as individual difference variance in the automatically activated evaluation of that attitude object. That is, for participants scoring about zero, it is inferred that there is no difference in automatically activated valence for blacks and whites. For participants having a positive priming score, it is inferred that pictures of blacks evoke a more negative evaluation in contrast to pictures of whites. However, what is taken for granted in this logic is that the procedure itself works for everyone. In other words, primes of *a priori* given universal valence (e.g., some items of the IAPS set) or individually selected positive and negative primes (see Fazio et al., 1986) would yield a robust affective priming effect for *everyone*. This, however, is not the case: Even if you successfully replicate the basic affective priming effect (i.e., if you use primes with universal positive and negative evaluation), you will get huge variance in the priming difference scores, with some participants even having a negative value. The same is true for semantic priming effects (see Stolz, Besner, & Carr, 2005). This error variance, of course, dilutes the measure of attitude-related evaluation (i.e., if you use, e.g., pictures of blacks and whites as primes) as well. If reliability of attitude-related priming scores is a severe issue, two routes need to be taken in future research: First, we should improve the basic procedure in such a way that it delivers a positive priming score for each participant if one uses clear positive and negative primes (in the affective priming paradigm) or if one uses clearly semantically related prime/target pairs (in the semantic priming paradigm). Variance around the mean should be low in this case. Note that in the ideal case (i.e., very low variance by using a set of nonambiguous positive and negative pictures) reliability of this variance is expected to be zero. If you have larger variance and this variance turns out to be reliable, there is the possibility of unwanted method variance (i.e., the method does not operate in the same way for each participant). The second route should be to check carefully (e.g., by simulation studies) whether RT-based measures have some intricacies that lower the applicability of classic test theory (see also, e.g., De Houwer, 2009).

Other recommendations. Because of space restrictions, we do not present detailed discussion of further points. We will, however, briefly mention them.

1. If you have a low number of trials per cell, a counterbalanced design (in which a given participant receives a given target only in one condition of the priming factor), and you expect only small effects (not unusual in priming research), you should consider increasing the power of your experiment by subtracting the main effect of balancing lists from error variance (see Pollatsek & Well, 1995).

2. Wellknown to psycholinguists is the long-standing issue of "the-language-as-a-fixed-effect fallacy," as introduced by Clark (1973). This often leads to the recommendation to add "by-item analyses" to the standard analyses. With regard to priming experiments, this is usually not necessary (see Raaijmakers, 2003; see also McNamara, 2005).

3. Modern statistical packages like SPSS are very convenient and offer many attractive opportunities. For example, the menu for the repeated measures analysis allows for inclusion of between-participants factors as well as inclusion of covariates. This might be taken as an invitation to reduce error variance to increase test power. If you are tempted to do so, please first consider Delaney and Maxwell (1981; especially p. 108) to prevent publication of nonsense.

Advanced data analytical methods

The present chapter does not allow for lengthy discussions of elaborate data analysis techniques. However, there are some interesting options that will certainly attract even more interest in the near future. Although we cannot present these methods of analysis in detail, we want to briefly mention them and draw attention to alternative ways of analyzing priming data

Diffusion model. As long as you use a binary decision task (e.g., lexical decision or evaluative decision) in your priming experiments, the application of Ratcliff's (e.g., 1978) *diffusion model* is a fascinating option. The model allows for separation of different components of the response process (i.e., decision threshold, bias, information uptake, and motor response; see Voss, Rothermund, & Voss, 2004; Voss & Voss, 2007, 2008).

Multinomial modeling. Another option might be the application of *multinomial modeling* (Riefer & Batchelder, 1988) to error data (e.g., Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Payne, 2001; Stahl & Degner, 2007). Depending on which processes you assume to determine a participants' performance in a priming task, you can design a model. For example, in the model by Payne (2001), the controlled processing of the target is separated from the automatic processing of the prime. Conrey and colleagues (2005) refined this model by separating different controlled and automatic processes; Stahl and Degner (2007) developed a model that included a further parameter to take guessing and response tendencies into account.

Hierarchical linear models. Basically, the structure of priming data is of hierarchical character: Trials are nested within participants. HLM's take the multilevel structure of repeated measures data into account (see, e.g., Locker, Hoffman, & Bovaird, 2007; Richter, 2006). Using them, you can model the influence of mixed predictor variables at the level of trials and participants as well as within-level and cross-level interactions. The advantage of these tools lies (1) in the straightforward application of regression analytical thinking to repeated measures experimental data and (2) the test of moderation of a within-participants effect by individual differences. HLM is the more advanced statistical technique compared with the older suggestions by Lorch and Myers (1990; see prior discussion).

Conclusions

Priming paradigms are a fascinating tool to study hypotheses of implicit social cognition. However, priming processes are a research domain at their own with a growing body of evidence that makes clear how complex the field is. We have tried to find a balance between giving an overview that informs about this complexity and giving practical advice for conducting priming experiments. Always remember that a priming result is, first of all, the outcome of an experimental procedure that might be open to different interpretations (see also Gawronski, LeBel, & Peters, 2007).

References

Abrams, R. L., & Greenwald, A. G. (2000). Parts outweigh the whole (word) in unconscious analysis of meaning. *Psychological Science, 11*, 118-124.

Bachmann, T., Luiga, l., & Pöder, E. (2005). Variations in backward masking with different masking stimuli: II. The effects of spatially quantised masks in the light of local contour interaction, interchannel inhibition, perceptual retouch, and substitution theories. *Perception, 34*, 139-154.

Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. *Psychological Science, 7*, 136-141.

Banse, R. (2001). Affective priming with liked and disliked persons: Prime visibility determines congruency and incongruency effects. *Cognition & Emotion, 15*, 501-520.

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: The Guilford Press.

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: Unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology, 32*, 104-128.

Becker, C. A., & Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 389-401.

Bessenoff, G. R., & Sherman, J. W. (2000). Automatic and controlled components of prejudice toward fat people: Evaluation versus stereotype activation. *Social Cognition, 18*, 329-353.

Bodner, G. E., & Masson, M. E. J. (2003). Beyond spreading activation: An influence of relatedness proportion on masked semantic priming. *Psychonomic Bulletin & Review, 10*, 645-652.

Breitmeyer, B. G., & Ögmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision (2nd ed.)*. New York, NY US: Oxford University Press.

Chan, E., Ybarra, O., & Schwarz, N. (2006). Reversing the affective congruency effect: The role of target word frequency of occurrence. *Journal of Experimental Social Psychology, 42*, 365-372.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12*, 335-359.

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

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating Multiple Processes in Implicit Social Cognition: The Quad Model of Implicit Task Performance. *Journal of Personality and Social Psychology, 89*, 469-487.

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology, 83*, 1314-1329.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 121*, 163-170.

De Houwer, J. (2009). Comparing measures of attitudes at the functional and procedural level: Analysis and implications. In R. H. F. R. Petty, & P. Brinol (Ed.), *Attitudes: Insights from the new implicit measures* (pp. 361-390). Mahwah: Erlbaum.

De Houwer, J., & Hermans, D. (1994). Differences in the affective processing of words and pictures. *Cogntion and Emotion, 8*, 1-20.

De Houwer, J., Hermans, D., Rothermund, K., & Wentura, D. (2002). Affective priming of semantic categorisation responses. *Cognition & Emotion, 16*, 643-666.

De Houwer, J., Hermans, D., & Spruyt, A. (2001). Affective priming of pronunciation responses: Effects of target degradation. *Journal of Experimental Social Psychology, 37*, 85-91.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. . *Psychological Bulletin, 135*, 347-368.

Degner, J. (2009). On the (un)controllability of affective priming: Strategic manipulation is feasible but can possibly be prevented. *Cognition & Emotion, 23*, 327-354.

Degner, J., & Wentura, D. (2009). Not everybody likes the thin and despises the fat: One's weight matters in the automatic activation of weight-related social evaluation. *Social Cognition, 27*, 202-220.

Degner, J., Wentura, D., Gniewosz, B., & Noack, P. (2007). Hostility-related prejudice against Turks in adolescents: Masked affective priming allows for a differentiation of automatic prejudice. *Basic and Applied Social Psychology, 29*, 245-256.

Dehaene, S., Naccache, L., Clec'H, L., G., Koechlin, E., Mueller, M., et al. (1998). Imaging unconscious semantic priming. *Nature, 395*, 597-600.

Delaney, H. D., & Maxwell, S. E. (1981). On using analysis of covariance in repeated measures design. *Multivariate Behavioral Research, 16*, 105.

Delplanque, S., N'Diaye, K., Scherer, K., & Grandjean, D. (2007). Spatial frequencies or emotional effects?: A systematic measure of spatial frequencies for IAPS pictures by a discrete wavelet analysis. *Journal of Neuroscience Methods, 165*, 144-150.

Deutsch, R., & Gawronski, B. (2009). When the method makes a difference: Antagonistic effects on 'automatic evaluations' as a function of task characteristics of the measure. *Journal of Experimental Social Psychology, 45*, 101-114.

Draine, S. C., & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General, 127*, 286-303.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*, 143-149.

Fazio, R. H. (1990). A practical guide to the use of response latency in social psychological research. In C. Hendrick & M. S. Clark (Eds.), *Review of personality and social psychology: Vol. 11. Research methods in personality and social psychology* (pp. 74-97). Newbury Park, CA: Sage.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *Journal of Personality and Social Psychology, 69*, 1013-1027.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*, 229-238.

Frings, C., & Wentura, D. (2003). Who is watching "Big Brother"? TV consumption predicted by masked affective Priming. *European Journal of Social Psychology, 33*, 779-791.

Frings, C., Wentura, D., & Holtz, M. (2007). Dysphorics cannot ignore unpleasant information. *Cognition and Emotion, 21*, 1525-1534.

Gawronski, B., Deutsch, R., LeBel, E. P., & Peters, K. R. (2008). Response interference as a mechanism underlying implicit measures: Some traps and gaps in the assessment of mental associations with experimental paradigms. *European Journal of Psychological Assessment, 24*, 218-225.

Gawronski, B., LeBel, E. P., & Peters, K. R. (2007). What do implicit measures tell us?: Scrutinizing the validity of three common assumptions. *Perspectives on Psychological Science, 2*, 181-193.

Glaser, J., & Banaji, M. R. (1999). When fair is foul and foul is fair: Reverse priming in automatic evaluation. *Journal of Personality and Social Psychology, 77*, 669-687.

Heil, M., Rolke, B., & Pecchinenda, A. (2004). Automatic semantic activation is no myth: Semantic context effects on the N400 in the letter-search task in the absence of response time effects. *Psychological Science, 15*, 852-857.

Hermans, D., De Houwer, J., & Eelen, P. (2001). A time course analysis of the affective priming effect. *Cognition & Emotion, 15*, 143-165.

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review, 10*, 785-813.

Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin, 96*, 29-44.

Kawakami, K., Dion, K. L., & Dovidio, J. F. (1999). Implicit stereotyping and prejudice and the primed Stroop task. *Swiss Journal of Psychology, 58*, 241-250.

Kawakami, K., Dovidio, J. F., Moll, J., Hermsen, S., & Russin, A. (2000). Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation. *Journal of Personality and Social Psychology, 78*, 871-888.

Klauer, K. C., Eder, A. B., Greenwald, A. G., & Abrams, R. L. (2007). Priming of semantic classifications by novel subliminal prime words. *Consciousness and Cognition, 16*, 63-83.

Klauer, K. C., Greenwald, A. G., & Draine, S. C. (1998). Correcting for Measurement Error in Detecting Unconscious Cognition: Comment on Draine and Greenwald (1998). *Journal of Experimental Psychology: General, 127*, 318-319.

Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and theories. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 7-49). Mahwah, NJ: Lawrence Erlbaum.

Klauer, K. C., Roßnagel, C., & Musch, J. (1997). List context effects in evaluative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 246-255.

Klauer, K. C., & Stern, E. (1992). How attitudes guide memory-based judgments: A two-process model. *Journal of Experimental Social Psychology, 28*, 186-206.

Klauer, K. C., & Teige-Mocigemba, S. (2007). Controllability and resource dependence in automatic evaluation. *Journal of Experimental Social Psychology, 43*, 648-655.

Klauer, K. C., & Voss, A. (2008). Effects of race on responses and response latencies in the weapon identification task: A test of six models. *Personality and Social Psychology Bulletin, 34*, 1124-1140.

Klinger, M. R., Burton, P. C., & Pitts, G. S. (2000). Mechanisms of unconscious priming: I. Response competition, not spreading activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 441-455.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*: Technical Report A-8. University of Florida, Gainesville, FL.

Larsen, R. J., Mercer, K. A., & Balota, D. A. (2006). Lexical characteristics of words used in emotional Stroop experiments. *Emotion, 6*, 62-72.

Locker, L., Jr., Hoffman, L., & Bovaird, J. A. (2007). On the use of multilevel modeling as an alternative to items analysis in psycholinguistic research. *Behavior Research Methods, 39*, 723-730.

Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology, 12*, 523-553.

Lorch, R. F., & Myers, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 149-157.

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review, 7*, 618-630.

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 3-23.

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 82-91.

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. New York: Psychology Press.

McNamara, T. P., & Holbrook, J. B. (2003). Semantic memory and priming. In A. F. Healy & R. W. Proctor (Eds.), *Experimental psychology* (pp. 447-474). New York: Wiley.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227-234.

Miller, J. (1988). A warning about median reaction time. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 539-543.

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132*, 297-326.

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology, 64*, 723-739.

Musch, J. (2000). Affektives Priming: Kongruenzeffekte bei der evaluativen Bewertung [Affective priming: Congruency effects in evaluative judgments]. *Unpublished doctoral dissertation, University of Bonn, Germany*.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106*, 226-254.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading. Visual word recognition* (pp. 264-336). Hillsdale, NJ: Erlbaum.

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1003-1019.

Olson, M. A., & Fazio, R. H. (2003). Relations between implicit measures of prejudice: What are we measuring? *Psychological Science, 14*, 636-639.

Otten, S., & Wentura, D. (1999). About the impact of automaticity in the Minimal Group Paradigm: Evidence from affective priming tasks. *European Journal of Social Psychology, 29*, 1049-1071.

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology, 81*, 181-192.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89*, 277-293.

Payne, B. K., Shimizu, Y., & Jacoby, L. L. (2005). Mental control and visual illusions: Toward explaining race-biased weapon misidentifications. *Journal of Experimental Social Psychology, 41*, 36-47.

Perea, M., & Gotor, A. (1997). Associative and semantic priming effects occur at very short stimulus-onset asynchronies in lexical decision and naming. *Cognition, 62*, 223-240.

Perea, M., & Rosa, E. (2002). The effects of associative and semantic priming in the lexical decision task. *Psychological Research, 66*, 180-194.

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 785-794.

Raaijmakers, J. G. W. (2003). A further look at the "language-as-fixed-effect fallacy". *Canadian Journal of Experimental Psychology, 57*, 141-151.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59-108.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114*, 510-532.

Ratcliff, R., & McKoon, G. (1978). Priming in item recognition: Evidence for the propositional structure of sentences. *Journal of Verbal Learning and Verbal Behavior, 17*, 403-417.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review, 95*, 385-408.

Richter, T. (2006). What Is Wrong With ANOVA and Multiple Regression? Analyzing Sentence Reading Times With Hierarchical Linear Models. *Discourse Processes, 41*, 221-250.

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review, 95*, 318-339.

Rothermund, K., & Wentura, D. (1998). Ein fairer Test für die Aktivationsausbreitungshypothese: affektives Priming in der Stroop-Aufgabe [An unbiased test of a spreading activation account of affective priming: Analysis of affective congruency effects in the Stroop task]. *Zeitschrift für Experimentelle Psychologie, 45*, 120-135.

Schmidt, T., & Vorberg, D. (2006). Criteria for unconscious cognition: Three types of dissociation. *Perception & Psychophysics, 68*, 489-504.

Shelton, J. R., & Martin, R. C. (1992). How semantic is automatic semantic priming? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 1191-1210.

Spruyt, A., Hermans, D., De Houwer, J., & Eelen, P. (2002). On the nature of the affective priming effect: Affective priming of naming responses. *Social Cognition, 20*, 227-256.

Spruyt, A., Hermans, D., De Houwer, J., Vandekerckhove, J., & Eelen, P. (2007). On the predictive validity of indirect attitude measures: Prediction of consumer choice behavior on the basis of affective priming in the picture-picture naming task. *Journal of Experimental Social Psychology, 43*, 599-610.

Spruyt, A., Hermans, D., De Houwer, J., Vandromme, H., & Eelen, P. (2007). On the nature of the affective priming effect: Effects of stimulus onset asynchrony and congruency proportion in naming and evaluative categorization. *Memory & Cognition, 35*, 95-106.

Stahl, C., & Degner, J. (2007). Assessing automatic activation of valence: A multinomial model of EAST performance. *Experimental Psychology, 54*, 99-112.

Stolz, J. A., & Besner, D. (1999). On the myth of automatic semantic activation in reading. *Current Directions in Psychological Science, 8*, 61-65.

Stolz, J. A., Besner, D., & Carr, T. H. (2005). Implications of measures of reliability for theories of priming: Activity in semantic memory is inherently noisy and uncoordinated. *Visual Cognition, 12*, 284-336.

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

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Van den Bussche, E., Van den Noortgate, W., & Reynvoet, B. (2009). Mechanisms of masked priming: A meta-analysis. *Psychological Bulletin, 135*, 452-477.

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition, 32*, 1206-1220.

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods, 39*, 767-775.

Voss, A., & Voss, J. (2008). A fast numerical algorithm for the estimation of diffusion model parameters. *Journal of Mathematical Psychology, 52*, 1-9.

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

Warren, R. E. (1974). Association, directionality, and stimulus encoding. *Journal of Experimental Psychology, 102*, 151-158.

Wentura, D. (2000). Dissociative affective and associative priming effects in the lexical decision task: Yes vs. no responses to word targets reveal evaluative judgment tendencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 456-469.

Wentura, D., & Brandtstädter, J. (2003). Age stereotypes in younger and older women: Analyses of accommodative shifts with a sentence-priming task. *Experimental Psychology, 50*, 16-26.

Wentura, D., & Degner, J. (in press). Automatic evaluation isn't that crude! Moderation of masked affective priming by type of valence. *Cognition & Emotion*.

Wentura, D., & Frings, C. (2005). Repeated masked category primes interfere with related exemplars: New evidence for negative semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 108-120.

Wentura, D., & Frings, C. (2008). Response-bound primes diminish affective priming in the naming task. *Cognition and Emotion, 22*, 374-384.

Wentura, D., & Greve, W. (2004). Who wants to be ... erudite? Everyone! Evidence for automatic adaptation of trait definitions. *Social Cognition, 22*, 30-53.

Wentura, D., & Greve, W. (2005). Assessing the structure of self-concept: evidence for self-defensive processes by using a sentence priming task. *Self and Identity, 4*, 193-211.

Wentura, D., Kulfanek, M., & Greve, W. (2005). Masked affective priming by name letters: Evidence for a correspondence of explicit and implicit self-esteem. *Journal of Experimental Social Psychology, 41*, 654-663.

Wentura, D., & Rothermund, K. (2003). The "meddling-in" of affective information: A general model of automatic evaluation effects. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in Cognition and Emotion* (pp. 51-86). Mahwah, NJ: Lawrence Erlbaum.

Williams, J. N. (1996). Is automatic priming semantic? *European Journal of Cognitive Psychology, 8*, 113-161.

Wittenbrink, B. (2007). Measuring Attitudes through Priming. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes.* (pp. 17-58). New York: Guilford Press.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology, 72*, 262-274.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Evaluative versus conceptual judgments in automatic stereotyping and prejudice. *Journal of Experimental Social Psychology, 37*, 244-252.

Author Note

Dirk Wentura, Department of Psychology, Saarland University, Saarbrücken, Germany, and Juliane Degner, University of Amsterdam, The Netherlands.

The research reported in this chapter was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG) to Dirk Wentura (WE 2284/3). Correspondence concerning this article should be addressed to Dirk Wentura, Department of Psychology, Saarland University, Building A2 4, P.O. Box 15 11 50, 66041 Saarbrücken, Germany, e-mail: wentura@mx.uni-saarland.de.

Table 1
Overview of tasks which dominate priming research

	Paradigm	Basic prime-target variation	Task	Basic references	Sample application in social psychology
Semantic priming					
	Semantic Priming	Semantic relationship, for example	Lexical decision	Neely (1991); McNamara (2005)	Wittenbrink et al. (1997)
		• Association	Pronunciation	Neely (1991); McNamara (2005)	Spruyt et al. (2007)
		• Category label (Prime), exemplar (target)	Semantic decision (e.g., object vs. person)	De Houwer et al. (2002); Klinger et al. (2000)	Banaji & Hardin (1996; Study 2)
		• category coordinates	Color naming	Warren (1972, 1974)	Kawakami et al. (1999)
	Semantic Priming (Masked versions)	*See above*	*See above*	Bodner & Mason (2003); Klinger et al. (2000); Wentura & Frings (2005)	Wittenbrink et al. (2001)
Response priming					
	Evaluative priming	Prime valence × target valence	(Valenced) Target evaluation	Fazio et al. (1986) Klauer & Musch (2003)	Fazio et al. (1995)
	Evaluative priming (Masked version)	Prime valence × target valence	(Valenced) Target evaluation	Draine & Greenwald (1998); Klauer et al. (2007)	Degner & Wentura (2009)
	AMP	Prime valence	(Neutral) Target evaluation	Payne et al. (2005); Murphy & Zajonc (1993)	Payne et al. (2005)
	Semantic category	Prime category × target category	Target categorization (e.g., gender)	Draine & Greenwald (1998); Klinger et al. (2000); Dehaene et al. (1998)	Banaji & Hardin (1996; Study 1)
	Weapon task	Black vs. white × weapon vs. no weapon	"Shoot!"/ "don't shoot"-decision	Payne (2001); Klauer & Voss (2008); Payne et al. (2005)	Correll et al. (2002)

Note: In the column "Basic references" we listed articles that either established a task, review evidence on this task, or report on recent studies using the task in basic research questions; the column "Sample application …" contains publication that apply the paradigm to social-psychological research questions concerning, e.g., stereotypes and prejudice.