Connectionist modelling of social judgement processes

Frank Siebler

This is the electronic version of my PhD thesis. The print version of the thesis does not have this cover page; the page numbers are however the same for both versions.

The APA-style reference to this document is:

Siebler, F. (2002). *Connectionist modelling of social judgement processes* [Electronic version]. Doctoral dissertation, University of Kent, Canterbury, UK.

The APA-style reference to the print version is:

Siebler, F. (2002). *Connectionist modelling of social judgement processes.* Doctoral dissertation, University of Kent, Canterbury, UK.

The research reported in Chapter 1 is formally published in:

Siebler, F. (2008). Emergent attributes in person perception: A comparative test of response time predictions. *Social Psychology, 39,* 83-89.

The connectionist model reported in Chapter 2 is also described in:

Bohner, G., Erb, H.-P., & Siebler, F. (2008). Information processing approaches to persuasion: Integrating assumptions from the dual- and single-processing perspectives. In W. B. Crano & R. Prislin (Eds.), *Attitudes and attitude change* (pp. 161-188). New York: Psychology Press.

Tromsø, August 25, 2009

Frank Siebler

Frank Siebler

PhD in Psychology

Connectionist modelling of social judgement processes

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the Faculty of Social Sciences at the University of Kent at Canterbury, January, 2002.

Abstract

The present work deals with classic and connectionist models of social judgement. Specifically, these approaches are compared in two domains: person perception and persuasive communication. In the domain of person perception, response time predictions are derived from three models of attribute emergence (Hastie, Schroeder, & Weber, 1990; Kunda, Miller, & Claire, 1990; Smith & DeCoster, 1998). Results from a laboratory study with a student sample suggest that classic and connectionist approaches to attribute emergence may complement each other. In the domain of persuasive communication, a constraint satisfaction model is developed that embodies an assumption common to three classic accounts of persuasion (ELM, Petty & Cacioppo, 1986; HSM, Chaiken, Liberman, & Eagly, 1989; Unimodel, Kruglanski & Thompson, 1999a). For both attitude judgements and cognitive response valence, results of simulation runs of actual studies show an excellent fit with human data as reported in the literature. The network model allows for the prediction of two consecutive attitude judgements of opposite valence to occur, within-subjects, under certain conditions. Therefore, first steps towards the development of an instrument for the repeated assessment of attitude judgements are undertaken. Specifically, the instrument shall allow to assess judgements both before and after they would be reported spontaneously. Results of two laboratory studies with student samples suggest that, with the experimental paradigm used, repeated assessment does not exert an undue influence on the second judgement. A predicted reversal in judgements' valence, in contrast, was not observed. Taken together, results showed that a connectionist approach to social judgement may complement classic approaches, may provide a viable alternative explanation of a domain's basic findings, and may stimulate the development of new research techniques.

Acknowledgements

Most of this research was conducted while I was a PhD student at the Department of Psychology, University of Kent at Canterbury. My thesis owes a lot to the stimulating research environment that I encountered at the Department, the insightful feedback that I received when presenting research ideas, and the ever available assistance in everyday matters. I would like to thank in particular Roberto Gonzáles, Anja Eller, and Charity Brown for their great support and steady encouragement in academic as well as other matters.

Next, I am indebted to the ALEKS research team at the University of Kassel, in particular Ernst-Dieter Lantermann, Ulrich von Papen, and Anja Springer. The colleagues in Kassel provided an outstandingly cooperative environment, a fact that helped extraordinarily with completing the write-up stage of my thesis.

Further, I am grateful for the invaluable contribution of my examiners, Lorne Hulbert, and Frank Van Overwalle, who gave excellent feedback on my thesis. I was more than happy to adopt their suggestions, and the thesis has certainly benefited considerably from these.

Then, I would like to thank my parents, Helga and Jürgen Siebler, who were always available for help with the many practical matters that arose during my stay abroad. Similarly, Monica Bär was an important source of encouragement and support.

Most importantly, however, I would like to thank Gerd Bohner who made all this possible. Beyond providing excellent academic advice, Gerd quite generally arranged for things to run as smoothly as they did. Definitely, supervision better than Gerd's is not conceivable to me.

Finally, many other people played an important role for the writing of this thesis, too many to mention all their names. I would therefore like to point out that I am grateful to all those who contributed remarks or suggestions on my research, or who made my time at the UKC the pleasant experience that it was.

Table of Contents

Introduction and overview

The present work deals with connectionist models of social judgement. Read, Vanman and Miller (1997) provided what may serve as a general and concise definition of the connectionist approach: "Processing in a connectionist model proceeds solely by the spread of activation among nodes, with the pattern of connections affecting how activation spreads. There is no higher order executive or control process. Moreover, knowledge in a connectionist model is represented entirely in the pattern of weights among nodes" (p. 28). More specifically, connectionist models comprise relatively simple processing units that conduct, in parallel, a basic information processing operation: they "sum the incoming activation following a specified equation, and then send the resulting activation to the nodes to which they are connected" (p. 27). Connections between nodes are weighted "where the weights affect how activation is spread" (p. 28). Thus, information processing in a connectionist model is processing of *numerical* information. In contrast to the connectionist approach's reliance on numerical computations, "... traditional assumptions about representation and process, which prevail throughout most of cognitive as well as social psychology, rest on a metaphor of the mind as *symbol* processor" (Smith, 1996, p. 894, italics added). Symbols (for instance, goals, traits, or stereotypes) are thought to be stored as discrete items; these items "are manipulated by rules that perform logical inferences and the like" (p. 894).

A connectionist approach to social judgement appears attractive for

various reasons. Compared to classic or symbolic approaches, connectionist explanations of given phenomena may at times be the more precise ones. For instance, with respect to schema- versus exemplar-based processing, Smith (1996) noted: "In traditional social psychological theories, the perceiver is said to search memory ... to locate the discrete schema or pattern that best fits the input information. This idea requires various assumptions, such as: How is the search performed ...? What are the criteria for stopping the search? When will a schema versus a well-known exemplar be used as the basis for inference? How are schemas formed in the first place?" (p. 900). With respect to a particular connectionist model of schema- or exemplar-based processing (McClelland & Rumelhart, 1986), Smith concluded: "In the connectionist model none of these questions arise. There is no search, and the learning process is explicitly modeled" (Smith, 1996, p. 900). In fact, connectionist accounts of psychological phenomena are usually specified in sufficient detail to simulate the phenomenon on the computer (for a recent discussion of the benefits of computer simulation for the development of a theory, see Mosler, Schwarz, Ammann, & Gutscher, 2001). Importantly, in order to answer questions or to allow for computer simulation, the connectionist paradigm does not add further assumptions to already existing symbolic accounts. Instead, as Fiedler (1996) stated (referring to his BIAS model), "... it is argued that the interaction of ecological principles with a few very basic notions about human information-processing capabilities can account for many phenomena that are traditionally understood in terms of motivations, goals, norms, or other categories of reasoned and planned action. ... No doubt those cognitive or motivational factors may be sufficient to influence judgments and behavior; however, it is strongly argued that they

are not necessary in many cases" (p. 193). Thus, connectionist explanations often provide quite parsimonious accounts of complex mental phenomena. Finally, various authors have pointed out that their connectionist models contributed to theoretical integration by "offering a single account for diverse findings that have typically been attributed to separate mechanisms" (Smith & DeCoster, 1998, p. 32; for related arguments, see Fiedler, 1996; Van Overwalle, Labiouse, & French, 2001).

More detailed discussions of similarities and differences between the connectionist approach and symbolic models are available in the literature. For an in-depth discussion with a focus on cognitive modeling see Smolensky (1988). The distinction of associative (connectionist) and rule-based (symbolic) processing has recently been discussed by Sloman (1996). Of great interest to social psychologists, Smith and DeCoster (2000) proposed a conceptual model that may integrate dual-process acounts from various domains of social cognition and cognitive psychology; their model draws on both associative and rule-based processing.

Within the connectionist paradigm, several approaches have been used. Thagard (1989) proposed a connectionist theory of explanatory coherence applicable to both scientific and everyday reasoning. In the networks used, nodes represent propositions. The coherence or incoherence between propositions is reflected in the weights of connections between nodes: whereas coherent propositions are linked by excitatory connections (i.e., connections with a positive weight), incoherent propositions share mutually inhibitory links (i.e., connections of a negative weight). Propositions may participate simultaneously in multiple relations of coherence or incoherence and, accordingly, have connections to more than one other

proposition. Competing theories or hypotheses are evaluated by spreading activation simultaneously along the connections. Specifically, one or more proposition nodes that correspond to observed evidence are activated externally. These nodes activate other nodes they share excitatory connections with, and deactivate other nodes they share inhibitory connections with. Thus, for a given proposition node, the actual activation value is constrained by the weights of the connections it shares, as well as by the activation values of the other nodes these connections are shared with. For each node, activation values are computed in each of repeated update cycles, until the network "settles" on a stable state where activation values change no longer. Now, the more strongly activated of competing theories or hypotheses may be determined. Thagard's network architecture has subsequently been employed by other authors to explain phenomena in the domains of, e.g., causal explanation (Read & Marcus-Newhall, 1993), goal and trait inference (Read & Miller, 1993), and impression formation (Kunda & Thagard, 1996). The architecture has, however been critisised for various reasons. The most prominent critique pertains to decisions that need to be made by a theorist when constructing a network: "which nodes (propositions) to include as relevant in a problem representation at all, what initial activation levels (belief strengths) to give the nodes, and what sign and weight to give each link" (Smith & DeCoster, 1998, p. 29). As alternatives, network architectures have been proposed that acquire these settings in a learning stage preceding the actual simulation of the target phenomena. For instance, Van Overwalle (1998) adopted a feedforward network to model causal explanation. The author's network developed connection weights from exposure to training data. Smith and DeCoster (1998) adopted a recurrent

network to simulate phenomena from person perception and stereotyping. The recurrent network did not only develop connection weights from training data, but also formed representations of concepts at that stage. These capabilities are certainly desirable because they do not only allow for the demonstration of a target phenomenon in a computer simulation, but also explain how the capability to produce the phenomenon may be acquired. On the other hand, the additional learning stage may come at the cost of a harder implementation into a computer model.

The following chapters will juxtapose classic and connectionist approaches in the domains of person perception, and of persuasive communication. Chapter 1 deals with a phenomenon from person perception. The phenomenon is the emergence of novel attributes, that is, attributes that are ascribed to members of a category combination although they are not ascribed to members of the constituent categories. I will oppose two classical explanations of the phenomenon (specifically, models proposed by Hastie, Schroeder, and Weber, 1990, and by Kunda, Miller, and Claire, 1990) with a connectionist account. Actually, two different connectionist accounts of the phenomenon can be found in the literature, one of them adopting hand-coded networks of concepts (e.g., Read & Miller, 1993; Read, Vanman, & Miller, 1997; see also Kunda & Thagard, 1996), and the other one using a network that acquires both concepts and connection weights in a learning procedure (Smith & DeCoster, 1998). I will focus on the weight-learning connectionist network model because that model allows for the derivation of predictions that differ pronouncedly from those of classical approaches. The results of an empirical study testing the predictions derived from classical models and the connectionist approach will be presented.

Chapter 2 is about persuasive communication. Here, I present a model of persuasion that is based on the hand-coded connectionist approach. I will show that findings of classical approaches to persuasion (namely the Elaboration Likelihood Model, Petty & Cacioppo, 1986; the Heuristic-Systematic Model, Chaiken, Liberman, & Eagly, 1989, and the Unimodel, Kruglanski & Thompson, 1999a) may also, and possibly more parsimoniously, be explained by that connectionist approach. More importantly, however, I go beyond most previous applications of the connectionist instrument in that I adopt a single network to simulate all of my target phenomena.

Chapter 3 describes steps towards the development of an assessment technique that allows for a most direct test of implications of the connectionist model discussed in Chapter 2. The results of two laboratory studies will be presented.

Chapter 1: Emergent attributes in person perception

The present chapter deals with a phenomenon from the domain of person perception. Research has shown that people, when describing members of multiple categories, sometimes use novel attributes - that is, attributes that are not used to describe members of the constituent categories. The emergence of novel attributes is not ubiquitous, but occurs more frequently with surprising or apparently conflicting category combinations than with others. I discuss three theoretical accounts of attribute emergence and present the results of a study designed to pinpoint the accounts against each other.

Introduction

Research on person perception has shown that perceivers, in order to make sense of apparently incongruent descriptions of a person, readily go beyond the information explicitly provided by the experimenter. Asch and Zukier (1984) for instance provided participants with person descriptions comprising a pair of attributes. For some of the pairs, the attributes were apparently antagonistic (e.g., "sociable" and "lonely"), whereas for other pairs, the attributes were congruent (e.g., "intelligent" and "witty"). The authors identified various strategies used by their participants to resolve apparent conflict between attributes. The modes of resolution revolved around a common theme: "The final product did not simply duplicate the initial information. The typical outcome was a sketch of a person (even if

schematic) or of an aspect of a person - an imaginative construction having a definite direction. In this manner two bare items of information were converted into the lineaments of a person" (p. 1239).

More recently, studies by Hastie, Schroeder, and Weber (1990), and by Kunda, Miller, and Claire (1990) have dealt with a related question. These authors provided participants with information about a target person's membership in social categories. Target persons were described as belonging to a single social category (e.g., "nurse"), to common category combinations (e.g., "female nurse"), to uncommon category combinations (e.g., "female mechanic"), or to combinations of categories with apparently conflicting implications (e.g., "blind marathon runner"). The authors found that members of uncommon or conflicting category combinations were ascribed novel or *emergent* attributes - that is, attributes that were used to describe a member of the category combination, but were not used to describe a member of each of the constituent categories alone (see below for the experimental procedures adopted). Similar to Ash and Zukier (1984), Hastie et al. as well as Kunda et al. discussed the emergence of novel attributes in terms of the perceiver starting with the information explicitly provided, and actively going beyond it in order to make sense of it. More pronouncedly than Ash and Zukier, both Hastie et al. and Kunda et al. outlined *general process models* of the particular cognitive activities that perceivers engage in when inferring novel attributes. I will detail the models' particulars below.

Connectionist explanations of the phenomenon have also been put forward in the literature (e.g., Read & Miller, 1993; Read, Vanman, & Miller, 1997; Kunda & Thagard, 1996; Smith & DeCoster, 1998). I will focus on Smith and DeCoster's (1998) account of attribute emergence because that

model allows for the derivation of predictions that differ pronouncedly from those of classical approaches. According to Smith and DeCoster's (1998) account of attribute emergence, novel attributes may not stem from a conscious effort that perceivers undertake in order to make sense of apparently conflicting information. Instead, they may stem from preconscious, automatic processes of conceptual combination. Again, details of the model will be provided below.

In the remainder of the present chapter, I will first describe the models proposed by Hastie et al. (1990), by Kunda et al. (1990), and by Smith and DeCoster (1998). From the models' respective processing mechanisms, I will derive empirically testable predictions that differ between models. Then, I present the results of a laboratory study testing these predictions.

"Emergent attributes": An empirical phenomenon and its three explanations

Assessment techniques

Attribute emergence has been demonstrated empirically in two different ways. Both of these techniques provide participants with labels of category combinations (e.g. "female mechanic"), or their constituents (e.g., "mechanic", or "female"), or both.

The first technique (used by Hastie et al., 1990, Study 1, and by Kunda et al., 1990, Studies 1 and 3) adopts an attribute-listing task. Specifically, participants are asked to describe, in an open-ended format, a person that belongs to the respective category or category combination. The

resulting protocols are later analyzed for attributes that participants employed in their description of the category combination (e.g., "female mechanic"), but not in their description of the respective single categories (here, "mechanic", and "female"). These attributes are considered emergent attributes.

The second technique (used by Hastie et al., Study 2, and by Kunda et al., Studies 2 and 4) does not require that participants generate attributes. Instead, attribute lists are provided by the experimenter. Participants' task is to rate how typical or descriptive these attributes are of members of single or combined categories. Attributes are considered emergent if they receive more extreme ratings for the category combination than for either of the constituent categories. Hastie et al. also use the term "outside rating" because with the rating task technique, an attribute qualifies as emergent if the rating for the combined category falls outside of the bounds defined by the ratings for the two constituent categories.

General outline of a study comparing the accounts

Below, I will describe three theoretical models that have been put forward as explanations of the emergence of novel attributes. Before doing so, I should outline how I am going to compare them subsequently in an empirical study. The approach is straightforward. With the assessment techniques described in the previous paragraphs, the field has an established research paradigm that may conveniently be re-used. As will be discussed shortly, all of the models rely on a different cognitive mechanism to explain the emergence of novel attributes. I will use these mechanisms to derive predictions from the models - predictions of response times. Consider a study

using the attribute rating task described above. Participants might be exposed to one of two category combinations: one combination shall be known to frequently elicit emergent attributes, the other combination shall be known to hardly ever elicit an emergent attribute. Are there reasons to assume that participants will respond more quickly in one of the conditions than in the other - and if so, why? To preview the outcome, each of the models provides a different answer to the question.

To sum up, my goal is to pinpoint three theoretical accounts of attribute emergence against each other. In order to do so, I will not introduce any new factor. Instead, I will collect a variable that has been available to, but has not been recorded in previous research: response times. Obviously, each of the models of attribute emergence can explain the emergence of novel attributes. The study to be reported below looks for the model that, in addition, predicts response times correctly. From the field's two established experimental paradigms, I will use the one that allows for greater experimental control of response times, namely the task of rating an experimenter-provided list of attributes for typicality of category members. In the next sections, I will derive each model's predictions of response times for that task.

Explanation 1: Causal reasoning triggered by surprise

Theoretical background

Similar to Ash and Zukier (1984), Kunda et al. (1990) assume that information about a person's membership in one or more categories leads to

the formation of a unified impression, or a set of expectations about that person. However, forming such an impression may be difficult at times - for instance when the constituents of a category combination have conflicting implications, or when category combinations are particularly surprising. Kunda et al.'s model of how persons resolve the conflict, or deal with the surprise, comprises the following sequence:

> "When confronted with a person who belongs to social categories with conflicting implications, people are surprised or puzzled. To resolve this puzzlement, they might ask themselves questions of the form: ... what might have caused a person belonging to one of the categories to acquire membership in the other? To answer such questions, people will ... construct an explanatory causal account of the reasons for the dual category membership. ... Guided by such causal accounts, one may include in the combined and unified image of the person some of the attributes associated with each of the original categories as well as some novel ... [attributes] ... that emerge from the causal reasoning used to resolve the puzzle of membership in both categories" (p. 552).

Thus, Kunda et al.'s (1990) model comprises four stages: a) experience of puzzlement, b) generation of a causal question, c) generation of a causal narrative, and d) construction of a unified impression. To explain the emergence of novel attributes, an additional assumption is made by Kunda et al.: namely that elements of the causal narrative may enter into, colour, or contaminate the unified impression that will ultimately be formed.

When combined with that additional assumption, the four-stage mechanism patently explains the emergence of novel attributes, as follows.

When exposed to *surprising category combinations*, people should be puzzled. In response to the puzzlement, they should generate a causal question. Trying to answer the question, they should next construct a causal narrative. The narrative, in turn, may comprise thoughts that the category combination's *constituents* whould not have elicited. Thus, when subsequently constructing a unified impression, the narrative may systematically colour the impression.

When exposed to *unsurprising category combinations*, or to a *single category*, people should not be puzzled to begin with. Accordingly, they should have no reason to generate causal questions, or to generate causal narratives. Because of having skipped over these three stages, people should not (at least not systematically) have thoughts unrelated to the constituent categories in mind when constructing a unified impression. Thus, no systematic contamination of the impression should occur.

In sum, according to Kunda et al.'s (1990) model, exposure to puzzling category combinations gives rise to causal reasoning. Causal reasoning does, however, not occur with non-puzzling category combinations or with single categories. Emergent attributes are the traces of causal reasoning found in people's unified impressions.

Response time predictions

What are the mechanism's implications for the time required to complete an attribute-rating task? Kunda et al. (1990) did not explicitly

address the issue of response latencies. However, their model is highly suggestive of a particular answer. First of all, however, it is important to note that, in a rating task, the processes described by the model should take place *before* the first attribute-rating is actually made. Recall that the model starts with a category combination being encountered, and ends with an impression being formed. In between, causal reasoning may or may not take place, and a narrative may or may not be constructed. Once the impression is formed, it may then be used to derive ratings of attributes for typicality or descriptiveness of category members. The point is important because it means that most of the cognitive work in the rating task is done before the first rating is actually made. If so, then what we should actually be interested in is how long it takes, under various conditions, to form an impression. In the following sections, I will adopt the point in time where people provide their first response as an indicator of how long it took them to form an impression. Thus, if people give their first rating relatively early (vs. late), I will infer that they required relatively little (vs. more) time to form an impression. Therefore, in the predictions to be derived now, I will assume a specific setup of experimental procedures, namely that response time scores of first ratings are assessed such that they *include* any time possibly required to form an impression. Similarly, I will assume the overall time required for doing the rating task (i.e., all of the ratings) to be assessed such that the score *includes* any time possibly required to form an impression.

Given the setup of experimental procedures just described, a prediction can be derived for the effect of causal reasoning on the overall time required to do the rating task. Kunda et al.'s model describes causal reasoning as a cognitive activity that does occur (puzzling combination

conditions) or does not occur (non-puzzling combination conditions, or single-category conditions) when forming an impression. Forming an impression via causal reasoning is certainly a non-trivial cognitive activity. In Kunda et al.'s (1990) model, it comprises four consecutive stages: First, puzzlement must be experienced. Then, a causal question must be posed. Next, a causal narrative must be constructed. Only now, after three preparatory stages, will an impression be formed. Even if the single stages involved in causal reasoning can be completed quickly, their aggregated time demands may add up to a noticeable delay when going through the whole sequence. Therefore, it appears no big leap to predict that participants who do engage in causal reasoning (puzzling-combination conditions) will require overall more time to complete the task than participants who do not engage in that extra activity (non-puzzling combination conditions, or single-category conditions).

The model allows for another, more specific prediction. Given a setup of experimental procedures as described above, people should generally require more time for their first ratings than for subsequent ratings. That is because, with the setup described, first-rating response times are assessed such that they include any time possibly required to form an impression. Of course, this effect of a rating's ordinal position (first vs. subsequent) stems from the setup of procedures. Over and above that theoretically uninteresting main effect, however, the prediction of a specific interaction can be derived from the model. Accordingly, any time delay that causal reasoning may cause should selectively affect first, but not subsequent ratings. That is because the first response is unlikely to be initiated until an impression is formed. Forming an impression, in turn, should take more time if causal

reasoning is involved (puzzling combinations) than if causal reasoning is not involved (unsurprising combinations). For subsequent ratings, in contrast, impression formation is not required because an impression has already been formed by then. To put the prediction differently: Response times of first ratings should, but response times of subsequent ratings should not differ systematically as a function of whether the category combination had initially triggered puzzlement or not.

Finally, the response time associated with a given rating should not differ as a function of whether the rating qualifies as an "outside rating" (indicating attribute emergence) or not. That is because in Kunda et al.'s model, all possibly time-consuming processes (e.g., puzzlement, generation of a causal explanation, etc.) have already been dealt with during impression formation. Therefore, whatever cognitive operation people adopt in order to assess how typical an attribute is of a category member, it may be the same for all attributes and yield "outside ratings" nevertheless.

Summary

To sum up, Kunda et al.'s model explains attribute emergence as a consequence of causal reasoning; causal reasoning, in turn, is triggered by the puzzlement experienced when category combinations either are particularly surprising, or have conflicting implications. If my analysis of the model is right, the model predicts A) that people should require more time overall to do the rating task when rating puzzling category combinations than when rating non-puzzling combinations or single categories; B) that the greater overall time required for puzzling combinations should be due to a

selective increase of response times of first (as opposed to subsequent) ratings; and C) that response times should not differ systematically between "outside ratings" and other ratings.

Explanation 2: Complex processes triggered by the failure of simple processes

Theoretical background

Hastie et al. (1990) propose that "subjects are doing something special to deal with the construction of an uncommon or novel conjunction" (i.e., category combination; p. 244). Specifically, with such conjunctions, they assume participants to engage in "more complex, deeper forms of reasoning about the conjunction" that result in "more complex types of explanations" (p 245). In their Discussion section, Hastie et al. outline a specific process model. The authors suggest that "a frame structure characterizes the long-term memory representation for generic categories" (p. 246). Frames are assumed to have "slots", some of which correspond to traits and attributes, for instance "gender, race, social class, personality attributes" (p. 246). For a given category, "each slot would be associated with a central tendency - a default value for the relevant attributes - as well as with an expression of the variability in admissible values of a member of the category" (p. 246). When exposed to a novel category combination, participants would need to create a new frame for the conjunction, and would need to fill the conjunction frame's attribute slots with appropriate values derived from the values found in the constituent categories' frames.

> "Given frame representations of simple categories, we believe that a two-stage model will be needed ... The first stage of this general model would proceed according to ... [relatively simple rules, for instance,] weighted averaging ... However, at some point during the application of a ... [relatively simple] process, it should be possible for that process to signal that it is 'in trouble,' perhaps because the default values expected on common attributes for the two ingredient-category frames are so discrepant that they signal a conflict or an impossibility. ... The signal from the first-stage processes that a difficult or surprising condition has been encountered would lead to a second stage that would involve the construction of a more complicated solution to the conjunction-attribute inference problem" (p. 246).

Similar to Kunda et al.'s (1990) approach, Hastie et al.'s model assumes some default processing of category information that may or may not be complemented by additional cognitive processes, depending on whether a given category combination is rare versus common (or, as they also call it, incongruent versus congruent). In Hastie et al.'s model, forming an impression of a category combination is described as the task of inferring, from the values found in the two constituent-category frames' slots, appropriate values for a new frame's attribute slots.

When dealing with *congruent category combinations*, an initial inference process based on simple rules (e.g., weighted averaging) should succeed. In this case, the new frame's attribute slots are being filled with values derived from, and similar to, one or both of its constituent categories.

Attribute emergence, as indicated by ratings that are less extreme or more extreme for the conjunction than for each of its constituents, is therefore unlikely.

When dealing with *incongruent category combinations*, an initial inference process based on simple rules may fail to find appropriate values to fill into the conjunction's attribute slots. Signalling the failure invokes a second processing stage that comprises inference processes following more complex rules. As the latter may draw on knowledge beyond the constituent categories (see below), the resulting attribute value may well be less or more extreme than the corresponding values of the constituents. Thus, when incongruent category combinations need to be dealt with by more complex processing, the emergence of novel attributes is likely.

How does the more complex inference process draw on knowledge beyond the constituent categories? Hastie et al. asked their participants to retrospectively provide explanations of their attribute ratings. From the taped protocols, the authors identified three general strategies used: a) the identification of a relevant case or exemplar in long-term memory; b) the application of general rules; c) a mental simulation of what kind of person it would take to handle that dual category membership.

It appears interesting to discuss strategy b) in some detail because, in my view, it may differ systematically from strategies a) and c). Both a) and c) seem to follow an overarching goal in that they attempt to identify (a) or construct (c) an exemplar that may represent the category combination. That appears useful because, once an exemplar is found, ratings for a broad range of attributes can be derived. Strategy b) seems to differ. Here, participants "rely on general rules abstracted from personal experience ...

these rules were activated to solve the problem of inferring the conjunction attributes. " (Hastie et al., 1990, p. 246). In fact, the examples provided by Hastie et al. seem to point to general heuristics that were applied in order to derive a rating for a single attribute: "For example: 'A woman in a man's job has to be tough...' Or, 'High achieving minorities tend to be very ambitious....' " (p. 246). I do not claim that my distinction of exemplar-based versus heuristics-based strategies were overly supported by this anecdotical evidence. However, the distinction opens the interesting possibility that some people may do most of the rating task with a solution applicable to many attributes already in their hands (i.e., an exemplar), whereas others go through the rating task looking for an applicable heuristic attribute-by-attribute. This may or may not actually be the case; the mere possibility will, however, become important when deriving response time predictions from the frame model of attribute emergence.

Response time predictions

What are the mechanism's implications for the time required to complete an attribute-rating task? Like Kunda et al. (1990), Hastie and colleagues (1990) did not explicitly address the issue of response latencies. Different from Kunda et al.'s model, Hastie et al.'s frame-based account of attribute emergence is highly suggestive of more than one answer. At the core of Hastie et al.'s processing mechanism is the need to fill the attribute slots of a newly created frame with appropriate values. When, exactly, are the values to fill in computed? Since Hastie and colleagues did not address the question, the frame model of attribute emergence (in its present

formulation) allows for several answers. First, we may assume that all of the values are computed and filled in when encountering a category combination. This view is similar to Kunda et al.'s model where a unified impression will be formed (with or without intermediate causal reasoning) as soon as category information is encountered, and before providing the first response. Hastie et al.'s model is compatible with that view, but is not committed to it. Alternatively, individual attribute values may be computed only when they are needed. Such a view would be incompatible with Kunda et al.'s model (where the goal is to form a unified impression), but is perfectly compatible with Hastie et al.'s model (where the goal is to have appropriate values in a newly created frame's attribute slots). Thus, the frame model allows in principle for the possibility that, when encountering a new category conjunction, people create a new frame, but do not fill in any values; instead, they may fill in the value for any given attribute only when they actually encounter the attribute in the rating task.

From the first version of the frame model (where attribute slots are filled in when encountering the category combination), the same predictions can be derived as from Kunda et al.'s model: A) uncommon or incongruent conjunctions (as opposed to common or congruent conjunctions) should trigger an additional processing stage and should therefore increase the overall time required to do the rating task; B) the greater overall time required for uncommon conjunctions should be due to a selective increase of response times of first (but not of subsequent) ratings, because all possibly time-consuming processes (i.e., second-stage complex processing) occur when initially filling in values into a newly created frame, not when subsequently rating it; and C) response times should not differ systematically

between "outside ratings" and other ratings (for the reason mentioned in B).

The second version of the frame model (where values are filled in only when they are needed) has partly different implications. We may still predict that the overall time required to do the task should be greater for uncommon or incongruent (as compared to common or congruent) conjunctions. However, if people compute appropriate values at the time they need them, the overall greater time is no longer due to a selective increase of response times of first ratings. Instead, increased response times may be found for ratings of attributes at arbitrary ordinal positions. Similarly, if people invoke second-stage complex processing not before, but during the rating task, then "outside ratings" should require more time than other ratings.

Summary

To sum up, Hastie et al.'s (1990) frame model explains the emergence of novel attributes for a category combination by the operation of a second processing stage. The second stage entails complex processes that may access knowledge beyond the information given in the constituent categories; it is invoked only if default processing in a first stage fails to infer appropriate values of the combinations' attributes from the constituents' values. If my analysis of the model is right, there are two versions of it that make partly different response time predictions. A) Both versions predict that people should require more time overall to do the rating task when rating rare or incongruent category combinations than when rating common or congruent combinations. B) The first version of the model points to prolonged first-rating response times as the cause of the overall greater time

requirement, whereas the second version identifies prolonged response times throughout the task as the cause. C) The first version predicts "outside ratings" to require the same amount of time as other ratings do, whereas the second version predicts "outside ratings" to take more time.

Explanation 3: Simultaneous application of multiple knowledge structures

Theoretical background

Smith and DeCoster (1998) reproduced several person perception phenomena in a recurrent connectionist network. Among these phenomena was the emergence of novel attributes. The network model uses associative processing (as opposed to rule-based processing; see, e.g., Sloman, 1996; Smith & DeCoster, 2000) as its cognitive operation. Different from Hastie et al.'s and Kunda et al.'s accounts of attribute emergence, Smith and DeCoster's (1998) model provides a mechanism whereby novel attributes need not be inferred by the perceiver after encountering a category combination. Instead, novel attributes may already have been added to the percept, by an automatic and preconscious process, when the combination enters a perceiver's consciousness. Smith and DeCoster (2000) describe how this may work:

> "Associative processing operates preconsciously ..., and we are generally aware only of the results of its processing. ... [Associative processing] generates what are experienced as intuitive and affective responses to objects or events. We may look at a mug and know that it

is used to hold coffee, or we may look at a friend and feel warmth and affection" (p. 111).

By the same mechanism, we may look at a member of a surprising or rare category combination and intuitively know that a given, novel attribute is typical of them. Smith and DeCoster (1998, Simulation 3) demonstrated the viability of an associative approach to attribute emergence by simulating a finding reported by Kunda et al. (1990) whereby people perceive a person who is both "Harvard-educated" and a "carpenter" as "non-materialistic". Because people do not ascribe non-materialism to persons described as either being "Harvard-educated" or being a "carpenter", "non-materialistic" qualifies as an emergent attribute.

In the following paragraphs, I will describe how Smith and DeCoster's (1998) model accounts for the emergence of novel attributes. In the interest of simplicity, I will introduce as few technical particulars of the specific network, or of associative processing more generally, as possible.

The associative network adopted by Smith and DeCoster (1998) is equipped with a cognitive operation called "pattern completion". Pattern completion is similar to drawing an inference: "Through training, an autoassociator learns predictive relationships among features of the inputs, and it uses this knowledge ... as it processes new stimuli. For example, a trained autoassociator exposed to an incomplete version of a stimulus pattern that it has previously processed will use what it has learned to fill in missing information" (p. 24). Importantly, the network is not limited to drawing one inference at a time, but can draw multiple inferences simultaneously.

To simulate the emergence of "non-materialism" when exposed to a

"Harvard-educated carpenter", Smith and DeCoster (1998) trained the network to represent three stereotypes. *Stereotype 1* may mean: "Harvard-educated persons are qualified for a high-paying occupation"; *stereotype 2* may mean: "Carpenters are low paid"; and *stereotype 3* may mean: "If a person is qualified for a high-paying occupation but is actually low paid, it may be because they are non-materialistic" (examples adapted from Smith & DeCoster, 1998, p. 29). Note that the stereotypes partially overlap: the concept of "being qualified for a high-paying occupation" appears in two stereotypes (1 and 3), and so does the concept of "being low paid" (stereotypes 2 and 3). Having trained the network to represent these knowledge structures, the authors then presented the network with a stimulus pattern indicating dual membership in the category combination "Harvard-educated person" and "carpenter", or, short, "Harvard-educated carpenter". As predicted, the authors found the network's output to include the pattern indicating "non-materialistic", that is, the emergent attribute. The explanation is straightforward. The network must have treated the "Harvard-educated" part of the actual input pattern "Harvard-educated carpenter" as an incomplete version of a pattern that it has processed before (during training), namely stereotype 1. Via pattern completion of stereotype 1, it must have inferred that the person is qualified for a high-paying occupation. Similarly, the network must have treated the "carpenter" part of the actual input pattern "Harvard-educated carpenter" as an incomplete version of a pattern that it has processed before (during training), namely stereotype 2. Thus, pattern completion of stereotype 2 led to the inference that the person must be low paid. Finally, the network must have treated the newly created bits of information ("the person is qualified for a high-paying occupation" and "the

person is low paid") as an incomplete version of yet another pattern that it has processed before, namely stereotype 3. By completing stereotype 3, the network inferred the emergent attribute, namely that the person is likely to be non-materialistic.

Key to the network's mechanism of attribute emergence is the partial overlap of the knowledge structures involved. Neither of the single categories "Harvard-educated" and "carpenter" is directly related to the emergent attribute "non-materialistic". Each of them is, however, indirectly related to the emergent attribute: first, via the stereotypes they participate in (stereotypes 1 and 2), and then, via another knowledge structure (stereotype 3) that allows for the inference of the emergent attribute. By drawing on all three knowledge structures, the network inferred the emergent attribute when presented with the category combination.

For Smith and DeCoster's (1998) purposes, simulating the exposure to category combinations that do not lead to the emergence of novel attributes was not required. To speculate briefly, such a condition might be simulated by training the network to represent stereotypes that have a lesser overlap. If, for instance, stereotype 2 ("Carpenters are low paid") was replaced with some other stereotype 2a (e.g., "Carpenters are rugged"), stereotypes 2a and 3 ("If a person is qualified for a high-paying occupation but is actually low paid, it may be because they are non-materialistic") would not overlap. As before, the network should treat the "Harvard-educated" part of the input pattern "Harvard-educated carpenter" as an incomplete version of a pattern that it has processed before (stereotype 1). Thus, pattern completion of stereotype 1 should lead to the inference that the person must be qualified for a high-paying occupation. Similarly, the network should treat the

"carpenter" part of "Harvard-educated carpenter" as an incomplete version of a pattern that it has processed before. This time, however, the previously processed pattern would be stereotype 2a. Thus, pattern completion of stereotype 2a should lead to the inference that the person must be rugged. In this scenario, the newly created bits of information ("the person is qualified for a high-paying occupation" and "the person is rugged") would not form the incomplete version of a pattern the network has processed before. Hence, no further inference would be drawn, and the novel attribute "non-materialistic" would not emerge.

Response time predictions

What are the mechanism's implications for the time required to complete an attribute-rating task? At first glance it may appear as if predictions were similar to those of the previously discussed models: inferring the emergent attribute requires drawing one inference more than not inferring it, so category combinations that elicit novel attributes should require more processing time than other category combinations. However, that is not the case. Smith and DeCoster (1998) emphasise that "the three stereotypes are not applied in a sequential fashion", but that "all simultaneously affect the processing of the new input stimulus" (p. 28). Similarly, describing pattern completion in terms of drawing inferences "does not imply that processing is sequential (i.e., first activate Stereotypes 1 and 2 and then Stereotype 3), for all stored knowledge ... actually affects processing simultaneously" (p. 29). Finally, as noted above, all of the inferences described are assumed to be drawn preconsciously using associative processing, "one striking feature [of

which] is how quickly and automatically it provides information" (Smith & DeCoster, 2000, p. 111). Thus, different from both other accounts discussed, the network model allows for the possibility that two category combinations may differ systematically in the number of novel attributes they elicit but, at the same time, may have A) no effect on the overall time required to do the attribute-rating task; B) no effect on the time required to rate first versus subsequent attributes; and C) no effect on the time required to provide an "outside rating" versus another rating.

Summary

To sum up, Smith and DeCoster's (1998) network model explains the emergence of novel attributes by means of an automatic, preconscious process. Depending on the overlap between previously learned knowledge structures, novel attributes may or may not emerge. If my analysis of the model is right, response times should not differ systematically as a function of whether category combinations elicit emergent attributes or not.

Study 1

The present study was designed as an extended replication of an experiment by Hastie, Schroeder, and Weber (1990, Study 2). Basically, I replicated their study, using parts of their materials and procedures, but assessing response times in addition.

Method

Participants

Forty-five undergraduate psychology students participated in partial fulfilment of course requirements. Thirty-two of these were female, 11 were male. Two persons chose not to disclose their gender. Participants' age ranged from 18 years to 38 years ($M = 22.64$).

Materials and procedure

General setup. Participants rated the typicality of attributes of members of various social categories. In order to assess response latencies, the study was conducted at the computer. Because a non-dichotomous response format (nine-point scales) was used, I selected the computer mouse as the input device.

Category labels. I adopted a subset of the category labels used by Hastie et al. (1990, p. 244). Specifically, participants were presented with the single category labels "female", "male", "nurse", and "mechanic", as well as with two combined category labels derived from the single categories. Depending on experimental condition, the combined categories were either "female nurse" and "male mechanic" (unsurprising combinations), or they were "male nurse" and "female mechanic" (more surprising combinations).

Trait descriptions. The list of 15 bipolar rating scales from Hastie et al. (1990, p. 244) was used. These scales are labelled at each end with opposing trait adjectives (e.g., "strong" - "weak"; see Table 1 for the scales). The adjectives of a pair were separated by response options consecutively

numbered from one (left-hand scale end) to nine (right-hand end of the scale). To select a response option, participants clicked on it with the computer mouse.

Table 1. *List of trait-adjective pairs used (Study 1).*

ambitious	unambitious
warm	cold
hostile	friendly
introverted	extroverted
intelligent	unintelligent
lower class	upper class
likeable	unlikeable
adventurous	cautious
honest	dishonest
calm	anxious
strong	weak
active	passive
dominant	submissive
imaginative	unimaginative
conscientious	careless

Note. *Trait-adjective pairs taken from Hastie, Schroeder, and Weber (1990, p. 244)*

Procedure. Participants were informed that the study was about person perception and impression formation. They learned that the study was

interested in their idea of the attributes that describe typical members of various groups. Using neutral example categories and example trait pairs (e.g., "police officer" and "trustworthy - untrustworthy"), the instructions explained how to respond on the nine-point scales. Next, participants were informed that their answers as well as their answer times would be recorded. They were encouraged to respond both quickly and accurately and further, not to take a break unless the programme indicated that it was appropriate to do so. At this stage, participants could choose to either review the instructions or to start the core of the experimental procedure.

The core procedure presented participants with a screen comprising an instruction, a category label, and one rating scale. The instruction appeared at the top of the screen. It was worded "Please click on that point of the scale that would best describe a typical ..." and was followed, on a new line, by a category label. The rating scale appeared, on a new line, below the category label. Selecting a scale point with the computer mouse triggered a one-second blank-screen interval. Then, a new response screen appeared that comprised the same instruction and category label as before, but displayed the next (of 15) response scales. The trait adjective scales were presented in the same, constant order as they appear in Table 1. For each response, two variables were recorded: firstly, the judgement, and further, the time an item had been visible on screen before a judgement was made. When the pool of bipolar adjective scales was exhausted for a given category label, participants were offered to either initiate the next step immediately, or to have a short break before doing so.

The sequence of 15 typicality judgements was then repeated using the same instruction and adjective scales, but replacing the category label by

another one. Seven category labels were used altogether. The first five labels were presented in a constant order: "Teacher" (a mere practice item), "Female", "Male", "Nurse", and "Mechanic". The two combined category labels were presented afterwards, in one of the two possible orders (see below for detail).

Participants were not forewarned about category labels. Instead, they encountered each label for the first time when presented with the first (of 15) descriptiveness judgement prompts for that label. Note that with this setup of experimental procedures (i.e., simultaneously displaying category label and response scale), response times included any extra time possibly required to process the category information. For the first rating of a category or conjunction, the category label was both new and unpredictable to participants; for subsequent ratings, the category label was both old and predictable to participants.

When having completed the rating task, participants responded to the manipulation checks. Specifically, the seven category labels (practice, single categories, and combined categories) were presented once more, one at a time. They were preceded by the prompt "How surprised would you be to learn that a person is (a) ..." and were followed by a nine-point response scale with the endpoints labeled "not at all surprised" (1) and "very surprised" (9).

Finally, participants reported their age, gender, and other possibly relevant variables.

Design and variables

Design and independent variables. Of focal interest were participants' responses to category conjunctions that comprised both a profession ("nurse" or "mechanic") and some background information ("female" or "male"). Category labels were constructed by prepending, to each of the two profession labels, either the unsurprising background information (unsurprising combinations: "female nurse", "male mechanic") or the more surprising background information (surprising combinations: "male nurse", "female mechanic"). Participants were exposed either to both of the surprising category combinations or to both of the unsurprising category combinations. To control for possible order effects, the combined category involving a (female or male) "nurse" was presented before the combined category involving a (female or male) "mechanic" for half of the participants, whereas the order of presentation was reversed for the other participants. Thus, the present study featured a fully factorial, mixed design with one within-subjects factor (target profession) and two between-subjects factors (surprisingness of target background, and presentation order).

Proportion of outside ratings. Participants' rating judgements on nine-point trait scales were recorded. For each category combination, a score indicating the proportion of novel attributes was computed following the procedure described in Hastie et al. (1990, p. 244). Specifically, a participant's ratings of a given trait scale were determined for the category combination as well as for the two constituent categories. If the rating for the combination was either more than one scale point higher than both or more than one scale point lower than both of the ratings for the constituents, the combination was credited one novel-attribute score. Novel-attribute scores

were determined for and accumulated over each of 15 trait scales. Then, scores were transformed into proportions. Thus, any category combination's score could range from 0 (no novel attribute) to 1 (all attributes novel).

Other dependent variables. Typicality judgements were assessed on nine-point scales. Response times were assessed with 100 ms precision. For the scoring of manipulation check variables, see the Procedure section above.

Results

Manipulation checks

I conducted a mixed-model ANOVA with participants' surprise ratings for combined categories ("[female/male] nurse" and "[male/female] mechanic") as the dependent variables. The within-subjects factor was target profession ("nurse" *vs.* "mechanic"); between-subjects factors were surprisingness of background information (target's gender: unsurprising *vs.* more surprising) and presentation order ("[female/male] nurse" presented first *vs.* "[male/female] mechanic" presented first).

As expected, participants expressed greater surprise when rating more surprising category combinations ($M = 5.65$) than when rating less surprising category combinations ($M = 2.57$), $F(1,41) = 35.41$, $p < .001$, $MSE = 6.02$. See Table 2 for cell means and standard deviations. In addition, I found a main effect of target profession whereby a "mechanic" triggered overall greater surprise ($M = 4.51$) as compared to a "nurse" ($M = 3.71$), $F(1,41) = 17.58$, $p < .001$, $MSE = .84$. Further, a two-way interaction between profession and surprisingness of background information occurred, as well as

a three-way interaction of all experimental factors, both *F*s(1,41) = 4.29, both *p*s < .05, *MSE* = .84. These interactions reflect that differences in the amount of surprise triggered by unsurprising versus more surprising category combinations were not of equal magnitude across levels of target profession and of presentation order. However, equal magnitudes of the differences are not required to consider the experimental conditions successfully established. As the pairwise comparisons depicted in Table 2 show, for each target profession in each presentation order, background information that was intended to trigger greater (*vs.* lower) surprise resulted in higher (*vs.* lower) surprise ratings, all *p*s < .005. - For other main effects and interactions, *F*s < 1.

To sum up, the manipulation check data were in line with expectations: category combinations intended to trigger greater surprise actually were rated as more surprising.

Table 2. Experienced surprise as a function of target profession, target gender, and presentation order (Study 1).

	Presentation order	
	Nurse, Mechanic	Mechanic, Nurse
Target profession		
Nurse		
Target gender		
female	2.55_a (1.63)	2.18_a (1.25)
male	5.00_b (2.52)	5.09_b (1.92)
Mechanic		
Target gender		
female	6.42_b (1.62)	6.09_b (1.81)
male	2.36_a (1.57)	3.18_a (2.14)

Note. Cell means depict level of surprise (1 = "not at all surprised", 9 = "very surprised"). Standard deviations are given in parentheses. Within professions and presentation orders, different subscripts indicate that means differ at $p < .005$. Cell N range from 11 to 12.

Emergent attributes, or "outside" ratings

From the 1350 category combination ratings available altogether (15 rating scales per combined category, times 2 combined categories per participant, times 45 participants), 68 ratings (5.04%) occurred outside the

bounds defined by the ratings for the constituent categories, plus or minus one scale point. They had been provided by 25 participants, whereas 20 participants did not generate an outside rating. How do the 5.04% outside ratings compare to the results of previous studies using the attribute rating task? Hastie et al. (1990) found a proportion of .28 outside ratings (their Study 2; p. 244). The proportion of outside ratings that I found is of a considerably lesser magnitude. It is, however, comparable in magnitude to the results of Kunda et al. (1990) who reported outside ratings in the magnitude of 9% (their Study 2; p. 563) and 3% (their Study 4; p. 569).

Participants' two outside rating scores each (pertaining to the conjunctions involving a "nurse" or a "mechanic", respectively) were submitted to a mixed-model ANOVA with the between-subjects factors "surprisingness of background information" and "presentation order", and repeated measures on "profession label".

For means and standard deviations, see Table 3. As expected, the analysis revealed a main effect of surprisingness of background information, such that a greater proportion of outside ratings was found when the background information was surprising ($M = .07$) than when it was not surprising ($M = .03$), $F(1,41) = 4.99$, $p < .04$, $MSE = .009$. A main effect of profession whereby the category "nurse" tended to elicit a somewhat greater proportion of outside ratings than the category "mechanic" did not reach marginal significance, $F < 2.4$, $p > .13$. For other effects, Fs < 1.

Table 3. Proportion of emergent attributes as a function of target profession, target gender, and presentation order (Study 1).

	Presentation order	
	Nurse, Mechanic	Mechanic, Nurse
Target profession		
Nurse		
Target gender		
female	.03 (.06)	.04 (.07)
male	.07 (.12)	.10 (.09)
Mechanic		
Target gender		
female	.07 (.09)	.06 (.06)
male	.02 (.04)	.02 (.03)

Note. Cell means depict proportions; standard deviations are given in parentheses. Cell N range from 11 to 12.

In sum, the overall proportion of outside ratings found was small, but comparable in magnitude to the proportions found in previous research using the attribute-rating task. As expected, outside ratings occurred more often in response to surprising (as opposed to unsurprising) category combinations. Further, the occurrence of outside ratings was not affected by other factors in the design. Thus, the empirical phenomenon "attribute emergence" has been

replicated in the present study. Therefore, it was appropriate to use the present data set in order to test response time predictions derived from the models of attribute emergence.

Response time data

Participants rated each of two category combinations on each of 15 trait-adjective scales. On average, they spent 2.78 s on a rating, the range was .71 s to 19.03 s. For analyses, response times greater than 6.00 s were recoded to 6.00 s (less than 4% of data points were affected)[1].

Response times for attribute scales

The two combined categories' response times for each of 15 trait-adjective scales were submitted to a mixed-model ANOVA. The between-subjects factors were surprisingness of target gender, and order of presentation. Repeated measures were on target profession ("nurse" *vs.* "mechanic"), and on ordinal position of the rating scale (1st to 15th).

I will first report effects that do not include the factor "ordinal position". These effects use not the single response times, but their mean across the

[1] The figure of six seconds was used as a (theoretically relatively arbitrary) cut-off criterion because this choice affected less than five percent of data points. The adoption of cut-off criteria, in turn, is a strategy frequently used in research adopting response times as dependent variables (see, e.g., Greenwald, McGhee, & Schwartz, 1998; Smith & Henry, 1996). This serves to avoid an undue influence of single, extremely large data points. Some researchers additionally log-transform response time data in order to further reduce the skewness typically found in response time distributions (e.g., Greenwald et al.). Recoding large scores to the upper limit is a more conservative strategy than deleting them. – A replication of the major response time analysis of Study 1 using the unrecoded scores led to the same conclusions as the results reported below.

15 attribute ratings pertaining to each category combination. Then, I will report effects that involve single response times, that is, response time per rating scale.

Effects involving aggregated response times

These effects test the three models' predictions of overall processing time requirements. To recapitulate predictions briefly: from Kunda et al.'s (1990) model as well as from both versions of Hastie et al.'s (1990) model I had derived the prediction that surprising (vs. unsurprising) category combinations should lead to greater (vs. lesser) overall processing time demands. From Smith and DeCoster's (1998) model, I had derived the prediction that processing time requirements should not differ between unsurprising versus surprising category combinations.

Cell means and standard deviations of average response times are depicted in Table 4. The time spent on the task may be computed by multiplying the means with 15. - The analysis revealed a main effect of surprisingness of target gender such that more surprising category combinations elicited greater response times (M = 2.93 s) than less surprising combinations (M = 2.47 s), $F(1,41) = 6.81$, $p < .02$, MSE = 10.20. This main effect was qualified by an interaction of surprisingness with target profession. A test of simple effects within levels of profession showed that surprisingness affected response times for a "mechanic" ($M_{surprising}$ = 3.02 s, $M_{unsurprising}$ = 2.40 s, $p < .003$), but not for a "nurse" ($M_{surprising}$ = 2.84 s, $M_{unsurprising}$ = 2.55 s, $p < .11$). For the interaction, $F(1,41) = 8.52$, $p < .01$, MSE = 1.04. Results for the "mechanic" are in line with predictions derived from Kunda et al.'s model and from both versions of Hastie et al.'s model.

Results for the "nurse" are in line with predictions derived from Smith and DeCoster's (1998) model.

Further, an interaction of target profession with order of presentation occurred. A test of simple effects within levels of presentation order revealed that, compared to a "nurse", a "mechanic" elicited greater reponse times if presented *before* a "nurse" ($M_{mechanic}$ = 2.90 s; M_{nurse} = 2.72 s; $p < .04$), but elicited marginally lesser response times if presented *after* a "nurse" ($M_{mechanic}$ = 2.52 s; M_{nurse} = 2.67 s; $p < .07$). For the interaction, $F(1,41) = 8.58$, $p < .01$, $MSE = 1.04$. The interaction was both unexpected and irrelevant for a test of the predictions derived from the models of attribute emergence.

Other effects involving aggregated response times did not occur, Fs < 2.2, ps > .14. For effects involving single response times, see the next section.

Table 4. Average response times of 15 attribute ratings as a function of target profession, target gender, and presentation order (Study 1).

	Presentation order	
	Nurse, Mechanic	Mechanic, Nurse
Target profession		
Nurse		
Target gender		
female	2.43_a (.52)	2.67 (.44)
male	2.91_b (.56)	2.77 (.77)
Mechanic		
Target gender		
female	2.84_c (.63)	3.19_c (.83)
male	2.20_d (.47)	2.60_d (.58)

Note. Cell means in units of seconds; standard deviations are given in parentheses.

a,b: Within professions and presentation orders, different subscripts indicate that means differ at p < .10. c,d: Within professions and presentation orders, different subscripts indicate that means differ at p < .05. Cell N range from 11 to 12.

Effects involving single response times

These effects test the three models' predictions of response times per

rating scale. To recapitulate predictions briefly: from Kunda et. al's model as well as from the first version of Hastie et al.'s model, I had derived the prediction that response times of first (but not of subsequent) ratings should be greater (vs. lesser) if conjunctions are more (vs. less) surprising. Thus, an interaction of ordinal position and surprisingness was predicted that should be due to the first rating. From the second version of Hastie et al.'s model, I had derived the prediction that response times of ratings at arbitrary ordinal positions should be greater (vs. lesser) if conjunctions are more (vs. less) surprising. The model would be supported by an interaction of surprisingness and ordinal position that is not due to the first rating, or by the absence of an interaction of these factors. From Smith and DeCoster's (1998) model, I had derived the prediction that response times should not differ systematically, at any level of ordinal position, between unsurprising versus surprising category combinations. The model would be supported by the absence of both a main effect of and interactions involving surprisingness.

Figure 1 depicts these data, separately for each level of target profession. The analysis showed a main effect whereby response times were strongly affected by a rating scale's ordinal position. Specifically, participants required more time for their first judgement (M = 4.55 s) than for any of the subsequent judgements (Ms in the range from 2.27 s to 2.83 s). Pairwise comparisons showed that the response time for the first rating differed from the response time for each of the remaining ratings, all ps < .001. This indicates that, as intended in the setup of experimental procedures, first-rating response times included the time required to process category information when first encountering it. For the scale position main effect, $F(14,574) = 28.77$, $p < .001$, $MSE = .89$.

Rating scale position was not qualified by a two-way interaction with surprisingness of target gender (counter to predictions derived from two of four models), $F < 1.1$, *n.s.* Scale position did however enter into a marginal three-way interaction with surprisingness of target gender and order of presentation, $F(14,574) = 1.64$, $p < .07$, $MSE = .89$. Thus, for one or more of 15 rating scales, response times tended to differ as a function of surprisingness and order of presentation. In order to see if the effect was due to response times of first ratings (as would be predicted by two of four models), I first conducted pairwise comparisons of the 15 response time scores within each level combination of target gender surprisingness (surprising vs. unsurprising) and presentation order ("nurse" vs. "mechanic" presented first). These showed, within each of four level combinations, that first-rating response time scores were greater than each of 14 other response time scores, ps $< .001$. This finding confirmed that response times had been assessed in sufficient precision to uncover systematic differences in the data if they exist. Then I compared, within each level of presentation order, the first-rating response time scores of surprising vs. less surprising conjunctions. These pairwise comparisons revealed that first-rating response time scores did not differ as a function of surprisingness, both ps $> .39$. Thus, the marginal three-way interaction was not due to systematic effects of surprisingness on first-rating response time scores. Other higher-order interactions involving both rating scale position and surprisingness of target gender did not emerge, Fs < 1. Therefore, the present data support Hastie et al.'s (second version) as well as Smith and DeCoster's (1998) models, but not Kunda et al.'s or Hastie et al.'s (first version) models.

Rating scale position further entered into an interaction with target

profession, $F(8.90,365.03) = 2.47$, $p < .02$, $MSE = 1.25$ (the Greenhouse-Geisser correction was applied because the interaction showed a significant W in Mauchley's test for sphericity). The interaction reflects that, for one or more of 15 rating scales, response times differed between the category combinations involving a "nurse" versus a "mechanic". The interaction was both unexpected and irrelevant for a test of the predictions derived from the models.

Finally, no other effects involving single response times were found, Fs < 1.1, ps > .40.

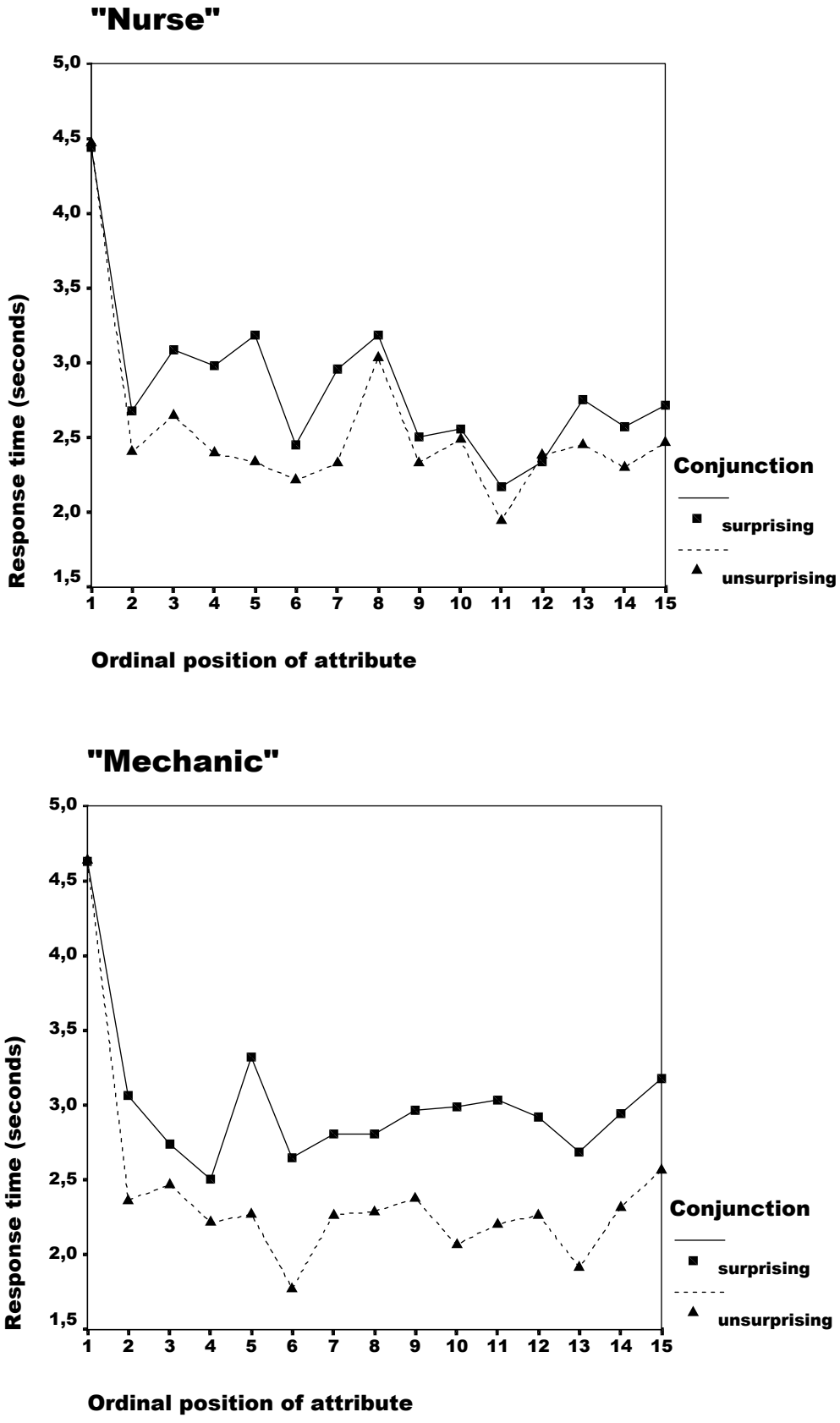

Figure 1. Response time as a function of target profession, conjunction surprisingness, and attribute position (Study 1).

Response time scores of outside vs. other ratings

As already mentioned, 68 (of 1350) conjunction ratings resulted in scores outside the bounds defined by ratings of the constituent categories, plus or minus one scale unit. The models of attribute emergence differed in their predictions of the time required to provide a rating score inside vs. outside of those bounds. To recapitulate: from Hastie et al.'s model (second version), I had derived the prediction that outside responses should take longer than other responses because they reflect complex second-stage processing that takes place at the time of the judgement. From Hastie et al.'s (first version) as well as Kunda et al.'s models, I had derived the prediction that outside ratings should not take longer than other responses because they reflect additional processing that has taken place when the conjunction was first encountered (i.e., before the rating task). From Smith and DeCoster's (1998) model, I had derived the prediction that outside ratings should not take longer than other responses because they reflect the result of preconscious, automatic processes.

In order to test these predictions, I conducted an analysis with response time scores as the unit of analysis (for a similar approach, see, e.g., Smith & Henry, 1996). Specifically, I submitted the 1350 response time scores to an analysis of variance with the factors target profession ("nurse" vs. "mechanic"), surprisingness of background information (target gender: surprising vs. not surprising), and rating score (inside vs. outside the bounds defined by the constituents, plus or minus one scale unit). This design resulted in single cell frequencies as low as seven; therefore, presentation order was not used as an additional factor.

For cell frequencies, cell means, and standard deviations, see Table 5. Results of the analysis need to be interpreted with caution, first because of one cell frequency being unusually low, second because of grossly inequal cell frequencies, and third because a test for variance homogeneity indicated inequal error variances, $F(7,1342) = 2.14$, $p < .04$, Levene test.

The analysis revealed a main effect of target gender surprisingness such that response time scores were greater when the target's gender was surprising ($M = 3.24$ s) than when it was unsurprising ($M = 2.52$ s), $F(1,1342) = 18.33$, $p < .001$, $MSE = 1.40$. Further, a main effect of rating score occurred. Accordingly, outside ratings took more time ($M = 3.09$ s) than other ratings ($M = 2.67$ s), $F(1,1342) = 6.07$, $p < .02$, MSE = 1.40. Next, target gender surprisingness and rating score tended to interact. Specifically, if target gender was surprising, outside ratings required more time ($M = 3.61$ s, $N = 50$) than other ratings ($M = 2.87$ s, $N = 640$), $p < .001$ for the pairwise comparison. If target gender was unsurprising, response time scores did not differ between outside ratings ($M = 2.56$ s, $N = 18$) and other ratings ($M = 2.47$ s, $N = 642$), $p > .70$ for the pairwise comparison. For the interaction, $F(1,1342) = 3.67$, $p < .06$, MSE = 1.40. Finally, for other effects, Fs < 1.5, ps > .20.

Reminding the reader that results need to be considered with caution (for reasons outlined above), findings may be interpreted tentatively as having the following implications. Hastie et al.'s model (second version) predicted greater latencies for outside ratings than for other ratings. This was actually found in the data if conjunctions were surprising. Smith and DeCoster's (1998) model predicted equal latencies for outside ratings as for other ratings. This was actually found in the data if conjunctions were

unsurprising. Thus, the present data support both of these models. Both Kunda et al.'s and Hastie et al.'s (first version) models predicted equal latencies for outside ratings as for other ratings. Although this was found in the data, it was found only with unsurprising conjunctions that, according to these models, should not have given rise to outside ratings to begin with. Thus, these models are not supported by the present data.

Table 5. Response latency as a function of target profession, target gender, and rating score (Study 1).

	Target gender	
	unsurprising	surprising
Target profession		
Nurse		
Outside rating		
no	2.54 (1.14)	2.78_{a} (1.15)
yes	2.78 (1.27)	3.59_{b} (1.45)
Mechanic		
Outside rating		
no	2.40 (1.17)	2.97_{a} (1.23)
yes	2.35 (.72)	3.64_{b} (1.42)

Note. Unit of analysis were individual response time scores. Cell means in units of seconds; standard deviations are given in parentheses. Cell N from left to right, from top to bottom: 319, 317, 11, 28, 323, 323, 7, 22.
a,b: Within professions and levels of surprisingness, different subscripts indicate that means differ at p < .02.

Discussion

The present study was designed to test predictions of response times derived from three theoretical models of attribute emergence. Methods,

materials, and procedures were adapted from an attribute-rating study by Hastie and colleagues (1990, Study 2). In addition to attribute rating scores, I recorded the associated response latencies.

The analysis of manipulation check variables showed that the experimental conditions had been established successfully: category combinations designed to trigger less (vs. more) surprise actually did trigger less (vs. more) surprise. The analysis of proportions of emergent attributes showed that the empirical phenomenon had been successfully replicated: the only significant effect in the data was a main effect of conjunction surprisingness such that more (vs. less) surprising category combinations led to greater (vs. lesser) proportions of emergent attributes. Thus, it was appropriate to test the models' response time predictions with the present dataset.

Effects involving aggregated response time scores

These effects test the models' predictions of overall processing time requirements. To recapitulate the relevant predictions: from both Kunda et al.'s (1990) and Hastie et al.'s (1990) model (both versions), I had derived the prediction of greater (vs. lesser) overall time demands if category combinations are surprising (vs. unsurprising). From Smith and DeCoster's (1998) model, I had derived the prediction that overall time demands should not differ as a function of category combination surprisingness.

The basic finding was an interaction of conjunction surprisingness with target profession. Accordingly, surprising (vs. unsurprising) category combinations triggered greater (vs. lesser) response latencies for a "mechanic", but not for a "nurse". These data provide support for each of the

models of attribute emergence. Response time results for a "mechanic" suggest that surprising conjunctions may, at least sometimes, increase the time required overall to do the rating task (in line with Kunda et al., and with Hastie et al., both versions). Response time results for a "nurse" suggest that surprising conjunctions may, at least sometimes, not increase the time required overall to do the rating task (in line with Smith and DeCoster, 1998). Recall that the analysis of outside ratings had revealed a significant and, importantly, not further qualified main effect such that greater surprisingness of conjunctions resulted in greater proportions of outside ratings. Together, these findings seem to suggest that two independent mechanisms of attribute emergence may exist: one mechanism that requires detectable amounts of processing time, and another one that works more efficiently. For additional analyses of response time scores, see Appendix A, where I discuss subjectively experienced surprise as a predictor of attribute emergence, and overall response time as a possible mediator of that effect.

Effects involving single response time scores

These effects test the three models' predictions of response times as a function of an attribute scale's ordinal position (first scale vs. subsequent scale). To recapitulate the relevant predictions, from Kunda et al.'s model, I had derived the prediction that puzzling combinations should selectively increase response times of first (but not of subsequent) ratings. Hastie et al.'s frame model allowed for two interpretations: either all attribute slots are being filled when first encountering the category combination (in which case predictions are the same as with Kunda et al.'s model), or, alternatively, attribute slots are being filled only when the attribute is actually to be rated (in

which case prolonged response times may be found for ratings of attributes at arbitrary ordinal positions). From Smith and DeCoster's (1998) model, I had derived the prediction that single response times should not reflect the level of surprisingness of the category combination.

Given the setup of experimental procedures, I expected first ratings to generally require more time than subsequent ratings - at both levels of conjunction surprisingness. Over and above taking generally more time, first ratings should (according to my analyses of Kunda et al.'s and Hastie et al.'s [first version] models) versus should not (according to my analysis of Hastie et al.'s [second version] and Smith and DeCoster's [1998] models) be further slowed down when conjunctions are particularly surprising. The only indication of a scale position by surprisingness interaction in the data was a marginally significant three-way interaction of these factors with presentation order. Within each level of presentation order, I found the expected scale position main effect whereby first responses generally take more time than subsequent responses. This confirms that the instruments used were, in principle, sufficiently sensitive to capture systematic variability in the data. With these instruments, I found first ratings' response latencies not to be affected by the surprisingness of a conjunction, at either level of presentation order. Thus, whereas people apparently do "something" time-consuming when encountering a new category combination, that "something" seems to be the same cognitive process for surprising and unsurprising conjunctions.

These results suggest that merely encountering a puzzling, rare, or uncommon category combination does not immediately give rise to processes likely to require noticeable amounts of processing time, like the construction of a causal narrative (Kunda et al.) or complex second-stage

processing (Hastie et al., first version). Instead, for uncommon category combinations, people may invoke extra processing stages only when required later on in the rating task (Hastie et al., second version), or may not need to invoke additional processes to begin with (Smith and DeCoster 1998).

Response time scores of outside vs. other ratings

This analysis dealt with the question if it takes more time to provide an outside rating than another rating. To recapitulate, from Hastie et al.'s model (second version) I had derived the prediction that outside ratings should take more time because they are computed at the moment of the judgement. From Kunda et al.'s as well as Hastie et al.'s (first version) models, I had derived the prediction that outside ratings should not require more time because they reflect the results of processes that take place when first encountering a conjunction, not when making a judgement. From Smith and DeCoster's (1998) model, I had derived the prediction that outside ratings should not require more time because they reflect the results of preconscious, automatic processing.

For surprising conjunctions, results showed that outside ratings take more time than other ratings. This finding lends support to Hastie et al.'s model (second version). According to the model, attribute values are computed when a judgement is to be made. For uncommon conjunctions, people may, at the moment of the judgement, have to invoke an additional processing stage involving complex operations. This mechanism simultaneously accounts for an outside rating's out-of-bounds response score as well as for its prolonged response latency.

For unsurprising conjunctions, results showed that outside ratings do not require more time than other ratings. This finding lends support to Smith and DeCoster's (1998) model. According to the model, outside ratings are the result of the automatic, preconscious combination of overlapping concepts. The mechanism simultaneously explains first, why outside ratings do occur for unsurprising conjunctions to begin with, and second, why these judgements are made relatively quickly.

The prediction of comparable response times for outside ratings as for other ratings had also been derived from Kunda et al.'s as well as Hastie et al.'s (first version) models. Both models explain attribute emergence by processing steps that take place when first encountering a category combination (not when making the judgement). However, these models predicted equal response times for conditions of *surprising* conjunctions, a prediction that was not met by the present data. Instead, the prediction was met in conditions featuring *unsurprising* conjunctions. In the latter conditions, the two models would not have predicted the occurrence of outside ratings to begin with. Thus, these models are not supported by the present data.

Summary and conclusions

The present chapter dealt with an empirical phenomenon from person perception, namely the emergence of novel attributes in descriptions of members of surprising category combinations. I described classical accounts of the phenomenon (Kunda et al., 1990; Hastie et al., 1990) as well as a more recent, connectionist explanation (Smith & DeCoster, 1998, Simulation 3). The classical accounts of attribute emergence revolve around

constructive processes that perceivers actively engage in when facing apparently incongruent conjunctions. The connectionist model, in contrast, invokes automatic, preconscious processes in order to account for the emergence of novel attributes.

From each model's cognitive mechanism, I derived predictions of processing time demands in the task of rating an experimenter-provided list of attributes for typicality of a member of the conjunction. Specifically, from each model's perspective, I predicted A) the time required overall to complete the task, B) the time required to initiate the first response, and C) the time required to generate an outside rating (as opposed to a rating inside the bounds defined by the constituent categories). Then, I tested these predictions in an empirical study.

The study adopted materials and procedures from the literature (Hastie et al., 1990, Study 2), but assessed response times in addition. Manipulation checks of surprise experienced confirmed that the experimental conditions had been successfully established. The actual proportions of novel attributes found in conditions of unsurprising versus surprising category conjunctions confirmed that the empirical phenomenon had been successfully replicated. Therefore, it was appropriate to test the predictions of response times with the present data set.

With respect to the overall time required to complete the task, the classical accounts predicted greater time demands for surprising than for unsurprising category combinations. The connectionist account, in contrast, predicted comparable time demands. An analysis of the overall time required to complete the rating task revealed that for one of the category combinations used (a male or female "mechanic"), data were in line with the classical

accounts. For the other conjunction (a female or male "nurse"), data supported the connectionist account. That response time results would differ between conjunctions had not been expected. More importantly, the proportions of emergent attributes had not shown a similar difference between stimulus conjunctions. Together, these findings suggest that people may be equipped with two distinct mechanisms whereby they infer novel attributes. One of the mechanisms may involve the kind of constructive, and probably time-consuming, processes described by classical approaches. The other mechanism may involve the preconscious, automatic combination of concepts described by the connectionist account.

With respect to the time required to initiate the first response, the connectionist account predicted comparable time demands for unsurprising and surprising conjunctions. The classical accounts assumed surprise to trigger additional processing steps that should occur either when encountering the conjunction (Kunda et al.; Hastie et al, first version) or when providing a judgement (Hastie et al., second version). Additional processing on encountering a conjunction should noticeably increase the time required to initiate the first response, whereas additional processing spread throughout the task should not. The analysis of first-rating response time scores provided no indication that surprisingness affected particularly first responses. Thus, explanations of attribute emergence that assume the perceiver to engage in demanding cognitive processes when facing puzzling category combinations (e.g., Kunda et al.'s construction of a causal narrative) are not supported by these data. Results are, however, in line with accounts that assume the perceiver not to engage in additional processing (as the connectionist model does), or to do so only later in the rating task (as Hastie

et al.'s model, second version, does).

With respect to the time required to provide a rating, the connectionist account predicted comparable response times for outside ratings as for ratings inside the bounds defined by the constituent categories. The second version of Hastie et al.'s frame model predicted greater time demands for outside ratings than for other ratings, whereas the other classical accounts predicted comparable time demands. Results showed that outside ratings of unsurprising conjunctions did not take more time than other ratings, whereas outside ratings of surprising conjunctions required more time than other ratings. When looking at response time results for surprising conjunctions, from the classical models, only Hastie et al.'s (second version) predicted response times correctly. The other classical accounts predicted response latencies that were actually found with unsurprising combinations - for unsurprising combinations, however, they expected outside ratings not to occur to begin with. Only the connectionist account provided an explanation for the occurrence of outside ratings when conjunctions are unsurprising (see below). With respect to unsurprising combinations, the connectionist model also predicted response times correctly.

Results for response latencies of outside ratings versus other ratings are, of course, of a correlational nature. Further, they need to be taken with caution because of partly low cell frequencies. Nevertheless, they fit nicely into the picture of two distinct mechanisms of attribute emergence outlined above. Accordingly, when rating attributes for typicality of *uncommon conjunctions*, perceivers may at times experience a conflict between the implications of the constituent categories. To resolve the conflict, they may engage in more complex processing as described by Hastie and colleagues.

That kind of processing, in turn, may lead to both an out-of-bounds rating score and a greater response latency. When rating attributes for typicality of *common conjunctions*, such conflict is less likely to be experienced. Novel attributes may nevertheless come about by automatic conceptual combination as described by Smith and DeCoster (1998). Specifically, if the conjunction (but not any of the constituent categories alone) overlaps sufficiently with a pre-existing knowledge structure (for instance, a stereotype), the overlap may trigger pattern completion of that structure. Because the connectionist mechanism draws on all knowledge structures simultaneously, inferring a novel attribute by means of preconscious conceptual combination does not imply a time penalty.

In sum, the present study's results suggest that a comprehensive account of attribute emergence will need to encompass both preconscious, automatic processing and conscious, time-demanding processing. Assessing participants' response latencies (in addition to the rating scores they provide) has proven a useful tool to disentangle the effects of preconscious and conscious cognitive processes that contribute to the inference of an emergent attribute. Future research may make further use of processing time requirements not only by recording response latencies, but also by experimentally manipulating the time available for a judgement. For instance, in response to unsurprising conjunctions, participants gave 18 outside ratings. If these ratings were actually due to automatic conceptual combination, they should be fairly stable against manipulations that reduce the time available to compute a judgement. In response to surprising combinations, participants gave a considerably greater number of outside ratings. If these ratings were actually due to complex second-stage

processing as described by Hastie and colleagues, limiting the available answer time is likely to interfere with the generation of outside ratings.

Having shown an application of the connectionist approach to the domain of person perception, we can now move on to exploring its usefulness in the domain of persuasive communication.

Chapter 2: Persuasive communication

In this chapter, I describe an experimental paradigm often used in research on persuasive communication, as well as the pattern of results typically found. Then, I outline how three current theoretical models of persuasion account for the typical findings. From a discussion of these models' shared assumptions, I derive an associative network of concepts that may also account for the result pattern. In Simulation 1, I show that the network model actually reproduces the pattern. In Simulation 2, I test the same network against the additional conditions, as well as the additional dependent variables, of an extended version of the experimental paradigm that has been used in the literature. In Simulation 3, I show that the network is rather insensitive to changes of the particular simulation parameters adopted.

Introduction

Persuasive communication aims to change people's attitudes by providing them with information. However, when trying to arrive at an attitude judgement, recipients of these influence attempts do not solely rely on the information explicitly provided. Instead, they take additional information into account. To mention just a few examples, attitude judgements have been shown to be impacted by people's prior expectations of message content (e.g., S. M. Smith & Petty, 1996), their desired processing outcomes (e.g., Giner-Sorolla & Chaiken, 1997), and their current mood (e.g., Bohner, Crow, Erb, & Schwarz, 1992) as well as other subjective experiences (e.g., Wänke,

Bless, & Biller, 1996). Although far from being comprehensive, this list of factors highlights that persons consider information from a wide range of sources unrelated to the persuasive appeal *per se*.

How do persons select from, possibly weight, and ultimately integrate the available information into an attitude judgement? The present chapter proposes the spreading of activation in networks of associated concepts as the cognitive operation that may underlie the selection and integration of information in many persuasion settings. Kunda and Thagard (1996) have recently taken a related approach in a different domain. These authors showed that simultaneous constraint satisfaction in spreading-activation networks qualifies as a plausible mechanism that may underlie the integration of stereotypes, traits, and behavioural information in impression formation. Specifically, the authors demonstrated in computer simulations that key phenomena from the impression formation literature can be modeled by means of activation spreading in connectionist networks. To show the viability of a spreading-activation account in the domain of persuasive communication, I adopted a similar strategy.

I first describe a classical study by Petty, Cacioppo, and Goldman (1981) and refer to other persuasion research with congenial findings. Then, I outline how three current models of persuasion may account for the general pattern of findings. These models are the Elaboration Likelihood Model (ELM; Petty & Cacioppo, 1986), the Heuristic-Systematic Model (HSM; Chaiken, Liberman, & Eagly, 1989), and the recently introduced unimodel (Kruglanski & Thompson, 1999a). From a discussion of these models' process assumptions, I derive a spreading-activation account of persuasion. To demonstrate the viability of my theoretical approach, I then simulate Petty,

Cacioppo, and Goldman's findings in a connectionist network as my first target phenomenon (Simulation 1).

A classic persuasion finding: Motivation and ability moderate the relative impact of cues and arguments on attitudes

Petty, Cacioppo, and Goldman (1981) provided their participants with a persuasive message advocating the institution of senior comprehensive exams at participants' own university. The authors varied three factors relevant to persuasion. First, *personal involvement* was manipulated by either telling participants that the exams might be introduced in the following year, or informing them that the exams might be introduced only ten years later. Thus, depending on the alleged time of introduction, participants would versus would not be affected personally by the proposed change. Second, *source expertise* was varied by ascribing the message to either a class at a local high-school (low-expertise conditions), or to a commission on higher education (high-expertise conditions). Finally, Petty et al. varied *argument quality* by constructing two message versions differing in the strength of the arguments used. Specifically, whereas all arguments were in favour of comprehensive exams, the "strong arguments were selected from a pool that elicited primarily favorable thoughts in a pretest, and the weak arguments were selected from a pool that elicited primarily counterarguments in a pretest" (p. 850). - Among other dependent variables, participants' post-message attitudes toward comprehensive exams, as well as toward the introduction of these exams were assessed. Because the two measures were highly correlated, they were combined into a single attitude index.

The results of Petty et al.'s study showed that the relative impact of source expertise and argument quality on attitudes differed between levels of personal involvement with the issue. Specifically, under conditions of low involvement, attitudes reflected source expertise, such that more positive attitudes were observed when the message allegedly stemmed from a source high (as opposed to: low) in expertise. Under conditions of high involvement, in contrast, attitudes reflected argument strength, such that more positive attitudes were observed when the message featured strong (as opposed to: weak) arguments. Thus, Petty et al. demonstrated that *personal involvement with the issue* is an important moderator of attitude judgements, determining whether attitudes are more strongly affected by context cues like source expertise, or, in contrast, are more strongly affected by the quality of the arguments put forward.

I consider the Petty et al.'s research a classical study because much of both preceding and subsequent research may be organised around their two by two by two factorial design. Specifically, many other studies may be described as adopting or extending the whole design or parts of it, and replicating findings with the same factors or conceptually. In particular the systematic variation of argument quality has become a ubiquitous manipulation in persuasion research, using either the same message topic (comprehensive exams - e.g., Petty & Cacioppo, 1984) or different ones (e.g., vitamin K intake, S. M. Smith & Petty, 1996, Study 2). Similarly, the manipulation of source characteristics has been replicated with other source-related cues (e.g., source credibility, Chaiken & Maheswaran, 1994), but comparable effects have also been shown with context cues that were not related to the source (e.g., the shere number of arguments, Petty &

Cacioppo, 1984, or the familiarity of the arguments' phrasing, Howard, 1997). Finally, the effects of issue involvement have been confirmed in subsequent research (e.g., Petty & Cacioppo, 1984), have been replaced by other motivational concerns (e.g., task importance, Chaiken & Maheswaran, 1994), by non-motivational manipulations (e.g., distraction introduced by a secondary task, Howard, 1997, Experiment 2), or by individual-difference variables (e.g., need for cognition, Howard, 1997, Experiment 3). Overall, these studies using either the same or conceptually equivalent variables confirmed the findings from Petty, Cacioppo, and Goldman's two by two by two design. Beyond the specific factors adopted by Petty et al., the results of all these studies may be summarised in a more abstract result pattern whereby attitude judgements reflect the evaluative implications of *message arguments* under conditions of high motivation and ability to process available information, but reflect the evaluative implications of *context cues* under conditions of low motivation or ability to process (see Table 6 for a schematic depiction).

Table 6. Schematic depiction of attitudes as a function of cue valence, argument quality, and processing motivation and capacity.

		Motivation / Capacity			
		low Arguments		**high** Arguments	
		weak	strong	weak	strong
Context cue	negative	−	−	−	+
	positive	+	+	−	+

Note. Minus- and plus-signs represent less versus more positive attitudes, respectively.

Three current theoretical models of persuasion

It has been noted that "the lion's share of current persuasion work was inspired by two major theoretical frameworks: Petty and Cacioppo's Elaboration Likelihood Model ... and Chaiken and Eagly's Heuristic Systematic Model" (Kruglanski & Thompson, 1999a, p. 84). As an alternative approach to persuasion, Kruglanski and Thompson proposed their Unimodel. Focusing on these three models' processing mechanisms, I will now briefly discuss how each of the models may account for the result pattern depicted in Table 6.

The Elaboration Likelihood Model

The Elaboration Likelihood Model (ELM, e.g., Petty & Cacioppo, 1986) distinguishes two routes to persuasion. The *central route* refers to attitude change as the result of "relatively extensive and effortful information-processing activity, aimed at scrutinizing and uncovering the central merits of the issue or advocacy", whereas the *peripheral route* refers to attitude change "based on a variety of ... processes that typically require less cognitive effort" (Petty & Wegener, 1999, p. 42). The distinction of central and peripheral routes describes the endpoints of an underlying continuum of elaboration likelihood. Accordingly, "the more motivated and able people are to assess the central merits of the attitude object ..., the more likely they are to effortfully scrutinize all available object-relevant information", whereas "at the low end of the elaboration continuum, information scrutiny is reduced" (Petty & Wegener, 1999, p. 42). Peripheral-route processing may differ from central-route processing *quantitatively*. That is, the same process (effortful scrutiny of available information) may be involved, but may occur to a lesser amount. Alternatively, peripheral processing may differ from central processing *qualitatively*, with attitude change stemming from non-elaborative processes " such as classical conditioning ..., self-perception ..., or the use of heuristics ..." (Petty & Wegener, 1994, p. 42). Central and peripheral processing may "co-occur and jointly influence judgments ... however, movement in either direction along the continuum will tend to enhance the *relative* impact of one or the other process ... on judgments" (Petty & Wegener, 1999, p. 59).

According to Greenwald's (1968) cognitive-response approach, the persuasive impact of a message is mediated by the valence of cognitive responses, or thoughts that persons generate in response to the message (for a recent test of the cognitive mediation hypothesis, see Romero, Agnew, & Insko, 1996). As a measure of elaboration or scrutiny (independent of attitudes as the to-be-explained variable), research in the ELM tradition has often assessed the amount and valence of issue-, message-, or argument-related thinking via thought-listing protocols (for detail, see Petty & Cacioppo, 1986, pp. 38-40). Actual scrutiny or elaboration of strong (weak) arguments should result in more positive (negative) thoughts. Put differently, "strong arguments should elicit more positivity in thought content than weak arguments, and hence greater message processing would lead to a greater difference between strong and weak arguments on ... [a] thought [valence] index" (S. M. Smith & Petty, 1996, p. 260). In addition to comparing the patterns of means of thought valence, the persuasive impact of elaboration has been demonstrated by showing that thought valence correlated more strongly with attitude judgements under conditions supportive to central-route (as opposed to: peripheral-route) processing, or that thought valence statistically mediated the impact of argument strength on attitudes (see S. M. Smith & Petty, 1996, for examples of both correlation and mediation analyses).

From an ELM point of view, the attitudinal pattern depicted in Table 6 reflects the result of effortful scrutiny of all available information, aiming to uncover the central merits of the issue, under conditions of high motivation and ability. Apparently, the strong (weak) arguments led to predominantly positive (negative) thoughts which, in turn, resulted in a favourable

(unfavourable) evaluation of the advocated position. Conditions of low motivation or ability, in contrast, triggered considerably less elaboration, if any. Instead, attitudes stemmed from more peripheral processing enhancing the relative impact of the context cue. This may be the case because the context cue was among the little information receiving any scrutiny at all (a quantitative difference to central-route processing), or because the judgement was derived by applying a heuristic involving the cue (a qualitative difference).

The Heuristic-Systematic Model

Like the ELM, the Heuristic-Systematic Model (HSM; e.g., Chaiken, Liberman, & Eagly, 1989) is a dual-process model. Two modes of processing are distinguished. In the *systematic mode*, "perceivers access and scrutinize all informational input for its relevance and importance to their judgment task, and integrate all useful information in forming their judgments" (Chaiken et al., p 212). In the *heuristic mode*, in contrast, "people focus on that subset of available information that enables them to use simple inferential rules, schemata, or cognitive heuristics to formulate their judgments and decisions", with an example of a persuasion heuristic being "Experts' statements can be trusted" (Chaiken et al., p 213). Further, the HSM assumes "that people must be motivated to process systematically and that this mode is adversely affected by variables ... that constrain people's capacities for in-depth information processing" (Chaiken et al., pp. 212-213). Because "heuristic processing [is regarded] as more exclusively theory-driven than systematic processing" and also because "the two modes can co-occur", the HSM's

"mode-of-processing distinction is not merely a quantitative one" (Chaiken et al., p 213).

As in the ELM, the valence of issue-, message-, or argument-related thinking is assumed to mediate argument strength effects on attitudes when both motivation and ability are high, but less so when either of these factors is low. Therefore, HSM research usually adopted thought protocols to gather a measure of systematic thinking (e.g., Chaiken & Maheswaran, 1994).

From an HSM perspective, the pattern of attitudes depicted in Table 6 points to attitude judgements being the result of heuristic processing under conditions of low motivation or ability. Accordingly, perceivers focused relatively exclusively on the context cue because the cue enabled them to use a simple persuasion heuristic in order to arrive at an attitude judgement. Heuristic processing may also have occurred under high motivation and capacity. Under these conditions, however, cue and arguments were further processed systematically, effortfully scrutinizing them, relating them to other knowledge that perceivers possessed, and considering their usefulness for the judgement task at hand. Apparently, when integrating all information into an attitude judgement, perceivers considered arguments to be more relevant, important, and useful than cue information.

The Unimodel

Different from both ELM and HSM, the unimodel (Kruglanski & Thompson, 1999a) does not distinguish qualitatively different routes to persuasion, or modes of processing. Instead, "the two persuasion types share a fundamental similarity in that both are mediated by if-then, or

syllogistic, reasoning leading from evidence to a conclusion" (Kruglanski & Thompson, 1999a, p. 90).

For instance (the examples in this paragraph are adapted from Kruglanski & Thompson, 1999a, p. 90), consider a persuasion setting where persons are exposed to a statement ascribed to Dr. Smith, a renowned environmental specialist, whereby "the use of freon in household appliances destroys the ozone layer, and therefore ought to be prohibited". If a person held the prior belief that "expert's opinions are valid" (major premise), and further realised that "Dr. Smith is an expert" (minor premise), he or she would likely conclude that "Dr. Smith's opinion (that the use of freon ought to be prohibited) is valid". Thus, syllogistic reasoning from persuasive evidence to a conclusion may explain persuasion by context cues. Importantly, the same mechanism (syllogistic reasoning) may also explain persuasion by arguments. Assume that a person held the prior belief that "anything that destroys the ozone layer should be prohibited" (major premise). With the persuasive communication providing evidence that "the use of freon ... does destroy the ozone layer" (minor premise), that recipient would likely conclude that "the use of freon should be prohibited". - Although there has been some debate about the appropriateness of equating persuasion by cues with peripheral or heuristic processing, and persuasion by arguments with central or systematic processing (cf. Bohner & Siebler, 1999; Chaiken, Duckworth, & Darke, 1999; Kruglanski & Thompson, 1999b; Petty, Wheeler, & Bizer, 1999), syllogistic reasoning appears to be a viable mechanism that may underly persuasion by each of these content types.

In line with dual-process models, the unimodel assumes an impact of motivation and ability on the processing of persuasive information.

Specifically, "if processing motivation and capacity are relatively low, only relatively simple and straightforward evidence will register" (Kruglanski & Thompson, 1999a, p. 90). Higher levels of motivation and ability are required "if the information is lengthy, complex, or unclear" because then more "cognitive work ... [is] involved in constructing the evidence from the various bits and pieces available to the recipient" (Kruglanski & Thompson, 1999a, p. 90). Interestingly, sometimes "the major premises that lend evidence its perceived relevance may need to be retrieved from memory" (Kruglanski & Thompson, 1999a, p. 90). These "memory search and activation processes ... in a proper sense ... constitute a 'cognitive response to persuasion' ... such activities often entail considerable 'cognitive work' that is quite painstaking and laborious" (Kruglanski & Thompson, 1999a, p. 90). Importantly, however, the amount of cognitive work required is in principle orthogonal to the content type (cue, or argument) of the evidence processed.

From a unimodel view, the attitude pattern depicted in Table 6 points to a confound of content type (cue versus argument) and presentation mode such that cues had probably been presented in a relatively simple and straightforward manner, whereas arguments had most likely been more lengthy, complex, or unclear. Under conditions of low motivation or ability, little cognitive work occurred. Therefore, only the more straightforwardly presented evidence (which happened to be the cues) registered, entered into syllogistic reasoning, and affected attitudinal conclusions. Under high motivation and ability, in contrast, participants did more cognitive work both collecting evidence from bits and pieces and activating relevant major premises in memory. Apparently, when that work had been done, the evidence hidden in the less straightforwardly presented arguments was found

to be greater, or was found to enter into syllogisms that were more relevant for the attitudinal conclusion.

Summary and conclusions

I described the process assumptions of three current approaches to persuasion, namely the Elaboration Likelihood Model, the Heuristic-Systematic Model, and the Unimodel. The models were found to differ with respect to the number of processes postulated, such that ELM and HSM assume two qualitatively distinct processes, whereas the unimodel postulates a single process. Between the dual-process models, in turn, differences were found with respect to the relation of the processes - whereas the ELM basically postulates a trade-off in the processes' relative impact on attitudes, the HSM emphasises that the modes of processing are orthogonal. Finally, the models differ in the specificity of the particular cognitive operations proposed. The dual-process models propose a few specific cognitive operations as underlying their low-effort process (e.g., classical conditioning, the application of heuristics, or schematic processing). With respect to cognitive operations underlying high-effort processing, they do not subscribe to any particular operation, but adopt relatively abstract descriptions of the process in terms of "scrutinising available information", "uncovering central merits", and "integrating useful information". The unimodel, in contrast, postulates a very specific cognitive operation (syllogistic reasoning) as underlying persuasion throughout.

Despite these differences, the models' explanations of the pattern of attitude judgements from Table 6 were similar in several respects. First, all

models' explanations of the effects of motivation and ability featured differences in the amount of processing involved. More specifically, for conditions of low (versus high) motivation and ability, the ELM predicted less (as opposed to: more) elaboration, the HSM predicted the processing of a subset of (as opposed to: all) available information, and the unimodel expected only the most straightforward (as opposed to: all) evidence to register. Put differently, there may or may not be *qualitative* differences between low-effort and high-effort processing - in any case, the three models seem to agree that there are (at least) *quantitative* differences. Second, all models' explanations of the effects of motivation and ability featured some kind of cognitive response emerging from high-effort processing (but not, or less so, from low-effort processing), and mediating the attitudinal impact of arguments. More specifically, for conditions of high (but not low) motivation or ability, the ELM expected to find argument elaborations (i.e., favourable or unfavourable thoughts); the HSM predicted the occurrence of cognitions concerning the relevance, importance, and usefulness of the informational input; the unimodel assumed the retrieval and activation of major premises that would otherwise not be active. Thus, the models seem to agree that self-generated thoughts are an integral part of high-effort processing.

I will now propose a spreading-activation account of persuasion. It is compatible with the two assumptions common to ELM, HSM, and unimodel - first, that differences between low- and high-effort processing include differences in the amount of processing, and second that, in high-effort processing, attitudes are mediated by cognitions that do not emerge in low-effort processing (see the next section for detail). Similar to the unimodel, the spreading-activation account features a very specific cognitive operation. As

a cognitive operation, the spreading of activation has two advantages over syllogistic reasoning: first, the operation is specified with sufficient precision to be implemented in a working computer programme, and second, the memory search and retrieval processes mentioned (but not specified in detail) by Kruglanski and Thompson (1999a) automatically fall out of the process - that is, no extra operations dealing with the search and retrieval of major premises are required (see the next section for detail).

Before introducing the account, I should note that I will use the terms "argument" and "cue" to represent concepts that have less versus more direct evaluative implications. Kruglanski and Thompson (1999a) showed that both cues and arguments may be set up such that their implications are easy versus difficult to extract. I acknowledge this possibility. For simplicity, I adopt the conventional notation nevertheless.

Associative networks of concepts, and persuasion

The pattern of attitude judgements described in Table 6 was the first target phenomenon that a spreading-activation model would have to account for. Building on assumptions shared by other current approaches, I now propose a spreading-activation account of persuasion and discuss how it may explain the pattern of attitude judgements. Having introduced my theoretical approach, I then describe its implementation in a connectionist network adopted to actually simulate the target phenomenon.

A spreading-activation model of persuasion.

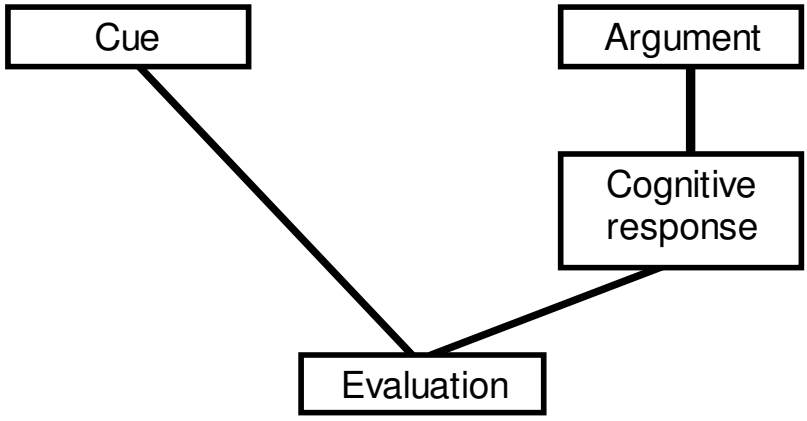

Figure 2. A generic, associative network of concepts. Boxes represent concepts relevant to persuasion, lines represent associations of concepts.

The above discussion of assumptions common to three models of persuasion translates readily into an associative network of concepts. In its most generic form (see Figure 2), the network distinguishes four categories of concepts relevant to persuasion: *cues*, *arguments*, *cognitive responses*, and *evaluations*. Further, the model assumes a particular pattern of associations between these concepts. Specifically, it assumes that cues are associated with evaluations in a relatively direct manner, whereas arguments are associated with evaluations more indirectly, via cognitive responses. This structure appears compatible with the other models' assumptions.

The *spreading of activation* along the associations is proposed as the cognitive process operating on the network. Accordingly, concepts in an active state should activate associated concepts. These, in turn, should pass activation on to more remote concepts. In the context of persuasion, cues

and arguments normally take the role of initially active concepts, whereas evaluations are typically found in the role of more remote concepts. Cognitive responses take the role of intermediate concepts mediating the spreading of activation between arguments and attitudes, but not (or: less so) between cues and attitudes.

As described in a previous section, both motivation and capacity to process available information have been found to exert a strong influence on attitude judgements. I do not actually model the mechanisms whereby motivation and capacity exert their impact. Instead, I assume that low motivation or capacity reduce the *amount of processing* devoted to the persuasive appeal.

How does the associative model account for the target pattern of attitude judgements whereby low-effort processing results in attitudes in line with cue valence, but high-effort processing leads to attitudes in line with argument quality (see Table 6)? If cues and arguments are presented simultaneously, I assume them to be processed in parallel. By virtue of a more direct link, evaluative concepts should receive activation *earlier* from cues than from arguments, allowing for a cue-based evaluative judgement in early stages of processing. Conversely, because of spreading via cognitive responses, activation from arguments should arrive at evaluative concepts relatively *late* and therefore impact in particular late stages of processing. However, the impact of arguments on evaluative judgements should be *greater* because it is the result of more complex thought involving a greater number of concepts overall. Cue information and argument information may sometimes have conflicting evaluative implications, for instance when a renowned expert delivers weak arguments. For such incongruent

combinations, the model predicts evaluative judgements of an opposite valence in early *versus* late stages of processing. To sum up, evaluative judgements from early stages of processing should reflect cue valence, whereas evaluative judgements from late stages should reflect argument quality. This pattern of predictions actually corresponds to the first target phenomenon.

The outlined network model depicts my theoretical approach. Before I proceed to test its viability by means of computer simulation, I should be more specific about several aspects of the model. First, the network described is assumed to be part of a greater network of concepts. Activation may spread to, or may come in from, other parts. For instance, in addition to having a direct impact on attitudes, cues may give rise to further cognitions. These further cognitions, in turn, might affect attitudes, resulting in an additional, indirect effect of cues on evaluations. For the present purpose, I do not consider such complex constellations, but focus on the basic assumptions that the effects of arguments on attitudes require mediation by other concepts, whereas cues exert their impact on attitudes directly. Second, although I just described the spreading of activation from cues and arguments towards attitudes, the model does allow for activation spreading along associations in both directions. And finally, I distinguished early *versus* late stages of processing. The distinction does not necessarily refer to physical time, but may represent any processing resource that can be expended to a lesser or greater amount (like, for instance, cognitive effort).

Implementation

Simulations were run on a 486 PC using a custom programme written in PASCAL. To actually simulate persuasion phenomena, I constructed a spreading-activation network based on those described by Thagard (1989), and Kunda and Thagard (1996). These networks consist of processing units representing social concepts, and of connections representing associations between concepts. *Processing units* can take on continuous positive or negative activation values (see below for the values' interpretation). The *connection* between a given pair of concepts is bi-directional and may be positive (indicating mutual excitation), negative (indicating mutual inhibition), or the concepts may be unassociated. To *simulate* a given phenomenon in a network, a relevant subset of its units is activated externally. Activation is then allowed to spread, along the connections, through the network. All active units spread activation both simultaneously and indiscriminately - that is, along all connections they have with other units. In the course of this process, units may change their activation state from an initially neutral level to a negative or positive level. As the spreading proceeds, the changes may become more pronounced for some units, and may reverse for others. Ultimately, the network settles on a stable state where units' activation levels do not change any more. Sign and magnitude of relevant units' activation levels are taken as the network's output. They are interpreted as indicating the likelihood that a person, in the situation modelled, would have the corresponding concept in mind. For instance, if a unit representing an evaluative concept showed a positive (negative) activation level, the network output would be interpreted as indicating a positive (negative) evaluation. A neutral activation level of the same unit, in contrast, would be interpreted as

indicating that the situation modelled did not have any particular evaluative implications[2].

Network setup. The network comprised seven units and six connections (see Figure 3). A special evidence unit (labeled "Observed" in Figure 3) was added. It merely served as a device of entering activation into the network. The network's topmost layer consisted of units representing the four kinds of information typically used in persuasion studies: One unit each represented a positive cue, a negative cue, a strong argument, and a weak argument. In simulation runs of the different experimental conditions, a relevant subset of these units (see below for detail) was linked to the special evidence unit, and activation was allowed to spread through the network. The activation level of the unit at the network's bottom layer was taken as the network's output, or attitude judgement.

[2] Thagard's (1989) network architecture focuses on changes of units' activation levels. The weights of connections between units, in contrast, are being held constant (see below for detail). Other architectures allow for a dynamic adaptation of connection weights during simulation runs (see, e.g., Smith & DeCoster, 1998; Van Overwalle, Labiouse, & French, 2001). As will become apparent shortly, the additional complexity introduced by using dynamic connection weights was not required to model my target phenomena.

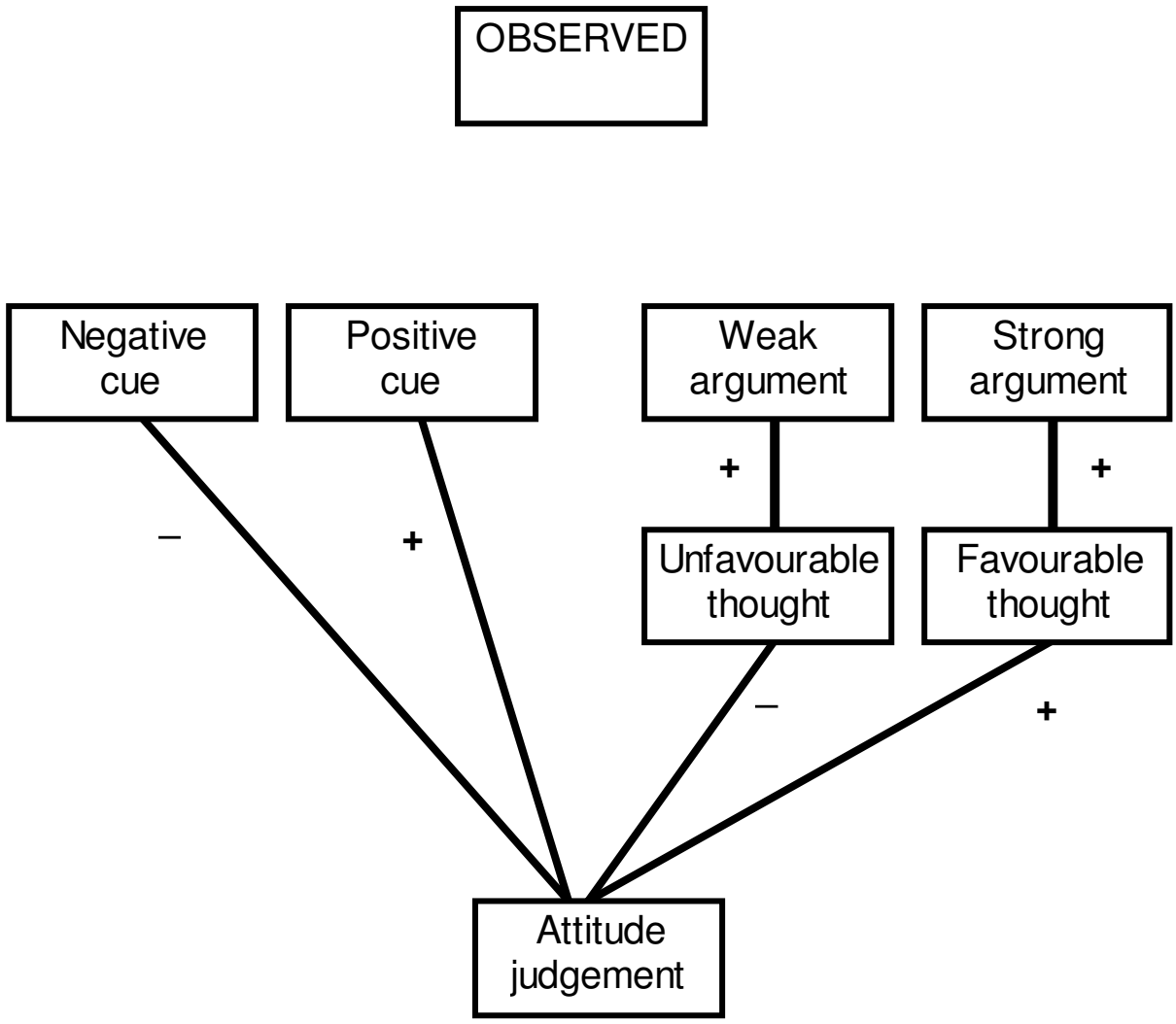

Figure 3. The parallel-constraint-satisfaction network used for simulating persuasion phenomena in Simulations 1 and 2. Boxes represent concepts, lines represent connections between concepts, plus- (minus-) signs depict mutual excitation (inhibition) of concepts.

As can be seen in Figure 3, the attitude judgement unit could receive positive or negative activation via two paths. First, it could receive activation from cue units, and second, it could receive activation from argument units. Importantly, argument units did not feed activation directly into the attitude judgement unit. Instead, an extra layer was added to this path, reflecting the assumption that persuasion by arguments (but not persuasion by cues) is

mediated by cognitive responses, or thoughts that are generated in response to the arguments provided. Note that the extra layer of cognitive responses introduced asymmetries with respect to both length and number of units involved into the otherwise symmetrical cue-attitude versus argument-attitude paths.

Simulation of experimental conditions. To simulate a given cue-argument combination, I connected the appropriate top-layer units to the special evidence unit. For instance, to simulate a condition where participants are presented with convincing arguments stemming from a highly credible source, I connected a) the unit labeled "strong argument", and b) the unit labeled "positive cue" (see Figure 3) to the special evidence unit. Activation was then allowed to spread, and activation levels were recorded. The final activation levels of relevant units were taken as the network's output under conditions of high motivation and capacity. In addition, activation levels from an early stage of processing were taken as network outputs under conditions of low motivation or capacity. In all simulation runs, I used outputs from the same, predetermined early stage (see below)[3].

[3] It is worth noting that the network model does not feature pre-stored links or associations between an attitude object and a summary evaluation (as, e.g., in Fazio's, 1995, model). Instead, attitude judgements are modelled as the result of dynamically integrating externally provided information with one's broader knowledge. Thus, in the present model, attitude *access* is a highly context-sensitive process of online computation that occurs, in parallel, across many conceptual nodes. Attitude *formation* may, in principle, be implemented by adding a mechanism for the adjustment of connection weights between conceptual nodes that participate in the computation of a judgement. As noted in Footnote 2, the present network architecture does not feature such a weight-adjusting mechanism. Finally, with its focus on the parallel spreading of activation as the mechanism whereby the online computation takes place, the model is best described as a model of *judgement formation*.

Simulation parameters. To compute each unit's level of activation in each of repeated update cycles, I adopted the nonlinear activation function from Kunda and Thagard (1996):

"On each cycle the activation of a unit j, a_j, is updated according to the following equation:

$$a_j(t+1) = a_j(t)(1-d) + \begin{cases} \text{net}_j(\max - a_j(t)) & \text{if net}_j > 0 \\ \text{net}_j(a_j(t) - \min) & \text{otherwise} \end{cases}$$

Here, d is a decay parameter (default = .05) that decrements each unit at every cycle, *min* is a minimum activation (-1), *max* is maximum activation (+1). Based on the weight w_{ij} between each unit i and j, one can calculate net_j, the net input to a unit, by:

$$\text{net}_j = \sum_i w_{ij} a_i(t)$$" (p. 308).

As a unit becomes activated more strongly (either positively or negatively), the function requires ever increasing amounts of activation by other units to activate it further into the same direction. The function also effectively restricts the possible activation values of units to a range from -1 to 1. Like Kunda and Thagard, I clamped the activation level of the *special evidence unit* to a value of 1. In Kunda and Thagard's implementation, the *connection weights* of excitatory versus inhibitory links differed in both sign and magnitude. For simplicity, my network's weights differed in sign only (but

see below). Specifically, I adopted weights of -.05 for inhibitory connections, and weights of .05 for excitatory connections.

Simulation parameters selected or changed on the basis of preliminary tests. Preliminary tests revealed that, with the present settings, activation spreading from cues dominated the activation level of the attitude judgement unit, effectively suppressing activation spreading from arguments. A more balanced setup would have required adding more argument and cognitive response units. Alternatively, it could be achieved by decreasing the magnitude of the connection weights between cue units and the attitude judgement unit. In the interest of simplicity, I chose the second option and used less strong weights for these links (positive cue: .025, negative cue: -.025). - As already mentioned, the spreading of activation through the network was simulated by repeatedly updating each unit's activation level. Simulations were run for 1000 update cycles. All units reached their final activation level (within three decimal positions precision) considerably sooner, and always before update cycle 400. In some conditions, two consecutive attitude judgements of opposite valence were expected. Read, Vanman, & Miller (1997) describe the operation of spreading activation or constraint satisfaction networks under conditions of unlimited processing time as follows: "... the activation of the nodes eventually reaches asymptotic values and the network stabilizes and stops changing" (p. 29) With the present network setup, the attitude unit was expected to ultimately reach an asymptotic value of a sign determined by the activated argument unit. Importantly, if the activated cue- versus argument-unit had opposite evaluative implications, the attitude unit was expected to reach an asymptotic value of a sign determined by the activated cue unit before. Preliminary tests

showed first, that the attitude unit did actually reach two consecutive asymptotic values of a different sign in these conditions (in the expected order - see Figure 4 for a graphical depiction) and second, that the first of these judgements reached its greatest magnitude in update cycle 24. For all conditions, I report activation levels from update cycle 24 as the network output representing low-effort processing, and activation levels from cycle 400 as the network output representing high-effort processing.

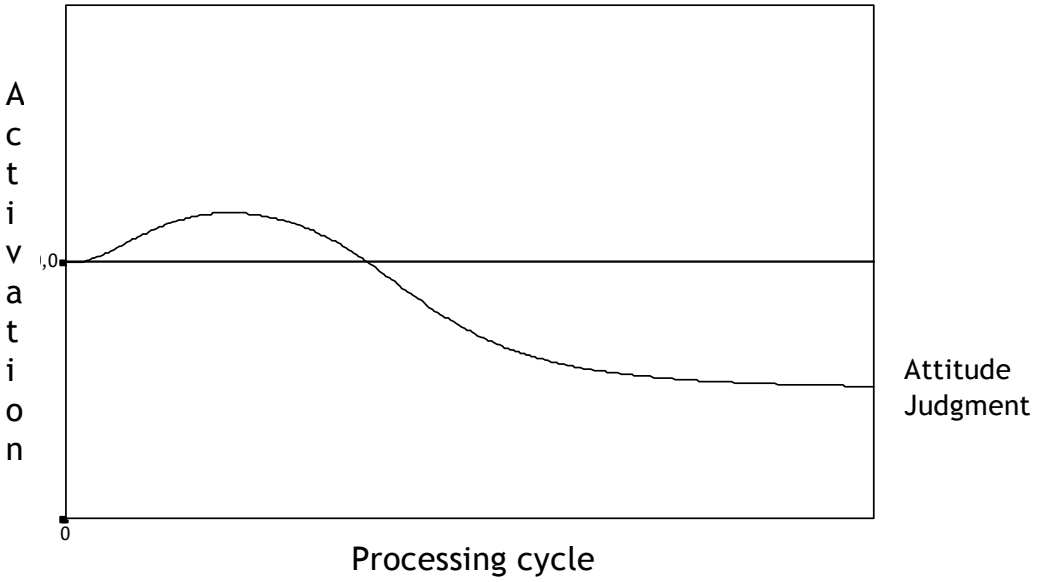

Figure 4. Activation history of the attitude unit as a function of processing cycle (here: from a simulation run of the "positive cue, weak argument" condition). The horizontal line indicates the initial activation of 0. As processing proceeds, the activation level reaches two asymptotic values opposite in sign.

The same network was used for the simulation of all conditions. That is, no units or connections were added or removed, except for the links to the special evidence unit required to simulate a given experimental condition.

Further, each connection weight's sign and magnitude was the same across simulation runs. Finally, before each simulation run, all units' activation levels were set to a neutral value of zero.

Simulation 1

Simulation 1 was conducted to show that the network does actually produce the predicted, qualitative pattern of attitude judgements (see Table 6). Specifically, in early stages of processing, the network should compute attitude judgements reflecting the evaluative implications of cues. Attitude judgements from late stages of processing, in contrast, should reflect the evaluative implications of arguments.

Of greater interest than the qualitative demonstration, however, was a quantitative comparison of network outputs with data from human participants. For the comparison, I chose the study by Petty, Cacioppo, and Goldman (1981) described previously. These authors experimentally manipulated source expertise (low *vs.* high), argument strength (weak *vs.* strong), and personal relevance of the issue (low *vs.* high). Petty et al. reported standardised attitude scores.

Simulation approach.

To simulate the four conditions of the core two (source expertise) by two (argument strength) design, I connected both one cue unit and one argument unit as appropriate to the network's special evidence unit, resulting in four independent simulation runs. In line with Petty, Cacioppo, and Goldman (1981), I considered low (high) expertise to be a negative (positive) cue. Further, like these authors, I assumed that the personal-relevance factor

had triggered low- vs. high-elaboration strategies in their participants. Accordingly, for each of the four simulation runs, I recorded the network's output at two stages of processing (24 and 400 update cycles, respectively), resulting in eight simulation scores altogether. Activation levels of the attitude judgement unit at the early (late) stage were mapped to conditions of low (high) personal relevance. As a measure of fit between simulation and human data, I computed the correlation between eight simulated attitude scores, and the eight standardised attitude scores reported by Petty and colleagues (for a similar strategy, see Van Overwalle, 1998). As a standard of comparison, I computed the fit between human data and the qualitative pattern of expected attitudes depicted in Table 6 (replacing plus- and minus-signs with 1 and -1, respectively).

Results.

Simulation. In early stages of processing, the sign of the attitude-judgement unit's activation mirrored cue valence (see Table 7, left panel). Specifically, independent of argument strength, that unit showed a negative activation in simulations of conditions involving a negative cue, but showed a positive activation when conditions involving a positive cue were simulated. In late stages of processing, in contrast, the sign of the attitude judgement unit's activation reflected argument quality (see Table 7, right panel). That is, independent of cue valence, the attitude unit showed a positive activation in simulations of conditions involving strong arguments, but showed a negative activation when conditions involving weak arguments were simulated. Finally, when the evaluative implications of cue and arguments were incongruent (that is, in simulations of "low expertise - strong arguments" and in "high

expertise - weak arguments" conditions), the attitude judgement unit consecutively showed activation levels of an opposite valence, in the expected order. Taken together, network outputs corresponded in sign to the expected data pattern.

Table 7. Activation level of the attitude judgement unit as a function of cue unit activated and argument unit activated, at two stages of processing (Simulation 1).

		Processing stage			
		early		**late**	
		Argument		Argument	
		weak	strong	weak	strong
Cue	negative	−.22	−.06	−.59	.45
	positive	.06	.22	−.45	.59

Note. Theoretical range of scores from -1 to 1. Greater scores indicate more positive attitudes. Early stage = 24 update cycles; late stage = 400 update cycles.

Comparison with human data. Simulation scores as well as the qualitative pattern of predictions (see Table 6) correlated strongly, and in the expected direction, with the attitude data reported by Petty, Cacioppo, and

Goldman (1981)[4], both *r*s(6) = .93, both *p*s < .002.

Discussion.

The network produced outputs in line with the expected result pattern, with early judgements reflecting the evaluative implications of cues, but late judgements being determined by argument quality. The predicted occurrence of two consecutive attitude judgements of opposite valence was observed. Furthermore, network outputs had an excellent[5] quantitative fit with data from human participants (as reported by Petty, Cacioppo, and Goldman, 1981), no worse than the prediction pattern from Table 6. Taken together, these findings lend support to the model and its process assumptions.

However, the model was developed with this result pattern in mind. Thus, the test may not have been a particularly strong one to begin with. To provide more conclusive evidence for the model's usefulness, I conducted another simulation and tested the network against a) a different set of attitude data from human participants, b) additional variables (beyond attitudes), and c) an additional persuasion phenomenon. Importantly, I used the same network, as well as the same simulation parameters, as before.

---

[4] In Petty, Cacioppo, and Goldman's (1981) study, participants in a non-factorial control group merely expressed their attitudes toward the message topic, to provide baseline data. As an approximation, I used the neutral activation level of zero and recomputed the measure of fit. Inclusion of the control group left results virtually unchanged, $r(7) = .97$, $p < .001$.

[5] For comparison, Van Overwalle (1998) reported correlation coefficients between simulation data and human data in the magnitude of .76 to .89.

Simulation 2

A discussion in the literature revolves around the interplay of the two processing strategies. In particular the HSM takes the pronounced position that the strategies are not mutually exclusive, but that they may co-occur and interact. Chaiken and Maheswaran (1994) reasoned that heuristic processing should also occur under conditions supportive to systematic strategies. If so, then heuristic processing could produce expectancies (for instance, about the probable validity of arguments) that might bias subsequent systematic processing. The biasing effect of heuristic on systematic processing should be particularly visible when persuasive communication is ambiguous, that is, amenable to differential interpretation.

To test their bias hypothesis, Chaiken and Maheswaran constructed (in addition to unambiguously weak or strong messages) ambiguous messages. These ambiguous messages comprised both weak *and* strong arguments. Biased systematic processing would be indicated by *high-effort, ambiguous-message* participants' attitude judgements reflecting the evaluative implications of *cue information*, whereas high-effort processing of unambiguous messages should result in attitudes reflecting argument quality (see Table 8 for the qualitative result pattern predicted by Chaiken & Maheswaran for the full design). In their study, the authors manipulated source credibility (low *vs.* high), message strength (unambiguously weak *vs.* ambiguous *vs.* unambiguously strong), and task importance (low *vs.* high). They reported attitude data as well as cognitive response data.

Table 8. Schematic depiction of attitudes as a function of cue valence, three levels of argument quality, and processing motivation and capacity.

		Motivation					
		low			**high**		
		Arguments			Arguments		
		weak	ambig.	strong	weak	ambig.	strong
Source credibility	low	−	−	−	−	−	+
	high	+	+	+	−	+	+

Note. Plus- and minus-signs represent positive and negative attitude scores, respectively. Ambig. = ambiguous.

Simulation 2 was designed as an extended replication of Simulation 1. I included additional ambiguous-message conditions to test for the occurrence of a data pattern indicative of biased systematic processing. Simulated attitude scores were compared qualitatively with the pattern depicted in Table 8, and were further compared quantitatively with human data as reported by Chaiken and Maheswaran. To allow for an interpretation of the fit between human and simulation data, I also computed the quantitative fit between human data and the pattern from Table 8 (replacing the pattern's plus- and minus-signs with 1 and -1, respectively), as a standard of comparison.

In addition, an index of thought valence was derived from simulation data (see below for detail) and was compared quantitatively to a measure of

systematic thought valence reported by Chaiken and Maheswaran (their "valenced index of attribute-related thoughts", p. 465)[6]. To allow for an interpretation of the fit between human and simulation data, I further derived a qualitative prediction pattern of thought valence from HSM assumptions and computed that pattern's fit with human data, as a standard of comparison. Specifically, under high motivation and ability, attitude judgements should be mediated by thought valence. Put differently, positive attitudes should stem from positive thought valence, and negative attitudes should be the result of negative thought valence. Therefore, to qualitatively predict thought valence under conditions of high motivation and ability, I used the same pattern that had been used for attitudes (as depicted in Table 8, right panel; plus- and minus-signs were replaced with 1 and -1, respectively). Under conditions of low motivation or ability, in contrast, systematic thinking should not occur to begin with. Therefore, to qualitatively predict thought valence under conditions of low motivation and ability, I used neutral scores of zero.

Simulation approach.

Attitude judgements. Technically, Chaiken and Maheswaran's (1994) study is a superset of the study by Petty, Cacioppo, and Goldman (1981), with eight of twelve conditions (i.e., conditions featuring unambiguous messages) being conceptually equivalent, and the remaining four conditions (i.e., conditions featuring ambiguous messages) being new. In line with

[6] Petty, Cacioppo, and Goldman (1981) assessed participants' recall of message arguments, but not (self-generated) cognitive responses. Since the network was not designed to

Chaiken and Maheswaran, I considered low (high) source credibility to be a negative (positive) cue. Further, like these authors, I considered task importance a factor that had influenced their participants' selection of processing strategies. Accordingly, network outputs from early (late) stages of processing were used as scores for conditions of low (high) task importance. With respect to conditions featuring unambiguous messages, I could thus re-use the previous simulation setup. Ambiguous message conditions were simulated by connecting the appropriate cue unit, as well as both the weak-argument *and* the strong-argument units, to the special evidence unit[7]. Given this setup, I expected the argument units to pass activation on to both cognitive-response units. The cognitive-response units, in turn, should pass activation on to the attitude-judgement unit. Importantly, the cognitive-response units' effects on the attitude unit should be of equal magnitude, but of an opposite sign. Therefore, they should cancel each other. Thus, ambiguous messages *per se* should have no net effect on the attitude unit. The attitude unit's activation level should instead be determined by cue valence, even under extended processing. Once activated, however, cue-consistent attitudes should pass activation on towards the cognitive-response units, thereby introducing an attitude-congruent bias into the latter units' otherwise balanced activation states. Note that this theoretical

---

simulate recall data, I could not compute a corresponding measure of fit in Simulation 1.

[7] With this setup, ambiguous *versus* unambiguous message conditions differed not only with respect to message ambiguity, but also in the number of units externally activated (three *vs.* two), and thereby in the total amount of activation introduced into the network. To control for possible effects, I additionally simulated no-message conditions where one cue unit, but *no* argument unit was activated externally. Between no-message and ambiguous-message conditions, corresponding network outputs on the attitude unit were virtually identical. Thus, the total amount of activation introduced into the network can be ruled out as an alternative

mechanism differs from the one proposed by Chaiken and Maheswaran. Specifically, Chaiken and Maheswaran hypothesised that, for ambiguous messages, "[source credibility] was expected to exert a direct, heuristic impact on attitudes under low task importance, and an indirect impact, mediated by biased systematic processing, under high task importance" (p. 462). The connectionist mechanism, in contrast, assumes that cues exert a direct impact on attitudes at both levels of task importance. More importantly, in this mechanism, biased systematic processing is a consequence (as opposed to: an antecedent) of attitude judgements.

Cognitive responses. Chaiken and Maheswaran (1994) computed a valenced index of systematic thought from thought-listing data by subtracting the number of listed thoughts rated as *negative evaluations* of the message object (a telephone answering machine) from the number of listed thoughts rated as *positive evaluations* of the object. For each simulated condition, I computed an analogous index by subtracting the activation level of the unit representing a *cognitive response to a weak argument* from the activation level of the unit representing a *cognitive response to a strong argument*. As with the simulated attitude data, separate indices were computed for early versus late stages of processing. The same update cycles as for simulated attitude scores were used (cycles 24 and 400, respectively).

Results.

Simulation runs of conditions featuring unambiguous messages naturally resulted in the same scores as in Simulation 1. These conditions will

---

explanation for attitudinal effects of message ambiguity in the network.

not be discussed in detail.

Simulation of attitude judgements. In early stages of processing, the sign of the attitude-judgement unit's activation reflected cue valence (see Table 9, left panel). Specifically, independent of argument strength and message ambiguity, that unit showed a negative activation when conditions involving a negative cue were simulated, but showed a positive activation in simulations of conditions involving a positive cue. In late stages of processing, in contrast, the sign of the attitude judgement unit's activation reflected argument quality when messages were unambiguously weak or strong, but reflected cue valence when the message was ambiguous (see Table 9, right panel). Specifically, for ambiguous messages, the attitude judgement unit showed a positive activation in the simulation of the condition involving a positive cue, but showed a negative activation when the condition involving a negative cue was simulated. Thus, the network produced the output pattern indicative of biased processing. Further, biased-processing attitude scores were of comparable magnitude as high-effort scores from simulations involving unambiguous messages.

Table 9. Activation level of the attitude judgement unit as a function of cue valence and three levels of argument quality at two stages of processing (Simulation 2).

		Processing stage					
		early			**late**		
		Argument(s)			Argument(s)		
		weak	ambig.	strong	weak	ambig.	strong
Context cue	negative	−.22	−.14	−.06	−.59	−.45	.45
	positive	.06	.14	.22	−.45	.45	.59

Comparison of attitude scores with human data. Unfortunately, Chaiken and Maheswaran (1994) did not report attitude scores for each of twelve cells of the full design. Instead, they reported attitude judgements within level of task importance, separately for each level of message strength, but collapsed over cue (i.e., six average judgements involving all twelve conditions of the design). Simulation scores correlated positively with these data from human participants, $r(4) = .98$, $p < .002$, and so did the qualitative prediction pattern, $r(4) = .94$, $p < .006$. For methodological reasons, Chaiken and Maheswaran had used two different renditions of the ambiguous message. They provided attitude data for each rendition, split by task importance and source credibility (i.e., four average judgements for each of two renditions of the ambiguous message). Again, simulation scores correlated positively with human data, $r(2) = .87$, *n.s.* for Chaiken and Maheswaran's message rendition 1, and $r(2) = .96$, $p < .05$ for their message rendition 2. The qualitative pattern fitted empirical findings for message

rendition 1 (but not for rendition 2) somewhat better than simulation scores did, $r(2) = .99$, $p < .01$ for rendition 1, and $r(2) = .96$, $p < .05$ for rendition 2.

Simulation of thought valence. I computed an index of thought favourability by subtracting, in each condition, the activation level of the unfavourable-thought unit from the activation level of the favourable-thought unit. In simulations of ambiguous message conditions, both of these units showed a positive activation at the early stage of processing, indicating that a person, when presented with both weak and strong arguments, would have both favourable and unfavourable thoughts in mind. However, the magnitude of activation was greater for the cognitive-response unit that was evaluatively congruent with the cue unit activated simultaneously. For instance, when the positive-cue unit was activated, the positive-response unit's score was .26, whereas the negative-response unit scored .19, indicating a somewhat greater likelihood of favourable than unfavourable thoughts. The asymmetry reflects a biasing impact of the activated cue unit, mediated by a long path of intermediate units (including the attitude unit). As processing continued, the activation level of the cognitive-response unit congruent to the activated cue increased to .51, whereas the activation level of the cognitive-response unit incongruent to the activated cue decreased to .06. Consequently, the difference score (or: thought favourability index - see Table 10) was of greater magnitude in late than in early stages. Thus, the biasing effect of cues on thought favourability increased with extended processing.

Table 10. Index score of thought valence as a function of cue valence and three levels of argument quality at two stages of processing (Simulation 2).

		Processing stage					
		early Argument(s)			**late** Argument(s)		
		weak	ambig.	strong	weak	ambig.	strong
Cue	negative	−.35	−.07	.18	−1.02	−.45	.94
	positive	−.18	.07	.35	−.94	.45	1.02

Comparison of thought valence with human data. Chaiken and Maheswaran reported the degree of thought favourability for each of the twelve conditions in their design[8]. These data from human participants correlated positively with the index of thought valence derived from simulation data, $r(10) = .97$, $p < .001$. The qualitative prediction pattern did not correlate quite as strongly, $r(10) = .89$, $p < .001$. The difference was significant, $t(9) = 2.61$, $p < .03$ for the test of difference between dependent *r*s.

Further results. In addition to the *pattern of means* of both attitude judgements and thought valence, the *correlation* between these variables has frequently been used in the literature as an indicator of processing effort. Specifically, a greater, positive correlation of attitudes and thought valence has been expected to occur under conditions conducive to high- (as opposed

[8] Chaiken and Maheswaran (1994) conducted path analyses to explore the mediational role of thought valence. Due to the small number of data-points available from the present simulation, corresponding analyses of network outputs could not be conducted.

to low-) effort processing (e.g., Smith & Petty, 1996). To see if the network model would produce a similar effect, I computed the correlation between simulated attitude scores and scores on the valenced thought index in the simulation data. Separate coefficients were computed for early versus late stages of processing; thus, each coefficient was based on six simulated scores. Both correlation coefficients indicated a positive association between attitudes and thought favourability. As expected, a greater correlation was found for late than for early stages (late stage: $r(4) = .97$, $p < .002$; early stage: $r(4) = .73$, *ns*). The difference was marginally significant, $z = 1.64$, $p < .06$, one-tailed, for the test of difference between independent *r*s from the same sample. Thus, over and above predicting appropriate *patterns of means* for attitudes as well as for thought valence, the network also properly reflected differences in the *association* of these variables at two stages of processing.

An analogous comparison could not be conducted for the qualitative prediction patterns, because the index of systematic thought valence should be a constant (0) under conditions of low task importance (where no systematic thinking is assumed to occur). To speculate briefly, if a correlation coefficient would be derived for these conditions by adopting additional assumptions, it would most likely be of a considerably lesser magnitude than the perfect correlation of 1.0 that the identical, qualitative prediction patterns for attitudes and thought valence under conditions of high task importance imply.

Discussion.

Simulation 2 was conducted as a more stringent test of the network model, using a different set of human data, an additional persuasion-relevant variable (cognitive responses), and an additional persuasion phenomenon (biased processing of ambiguous messages), but the same network and network configuration as Simulation 1. Network outputs on relevant units were not only qualitatively in line with predictions, but furthermore showed good to excellent fit with Chaiken and Maheswaran's (1994) data from human participants. For attitude judgements, I found a high correlation of human data and simulation data both across all conditions and within the newly introduced ambiguous-message conditions. It was comparable in magnitude to a qualitative pattern of predictions derived from these authors' theoretical considerations. Interestingly, for thought favourability, network outputs correlated even more strongly with the laboratory data reported by Chaiken and Maheswaran than a qualitative prediction pattern derived from their theoretical approach. Finally, the association of attitudes and thoughts within the simulation data differed between stages of processing in the expected way. Taken together, these findings lend strong support to the network model and the underlying process assumptions.

Van Overwalle, Labiouse, & French (2001) adopted a recurrent connectionist network to model, among many other phenomena from social cognition, Chaiken and Maheswaran's (1994) study on biased processing. Their network employed a cognitive mechanism different from the mere spreading-activation account used in my simulations. Specifically, in Van Overwalle et al.'s network, connection weights automatically adapt during the course of a simulation run. Thereby, the authors achieve a permanent

change to connection weights, an effect that conveniently models the temporal stability of central-route attitude change that is typically found in human participants. On the other hand, their model does not seem to provide simulation data on cognitive responses. Thus, the cognitive mediation of arguments' effects on attitudes (as the theoretical mechanism underlying central-route attitude change) appears not to be part of the model. In sum, each of the two network models accounting for the Chaiken and Maheswaran results has its specific strengths and limitations.

Results from both Simulations 1 and 2 spoke to the viability of a spreading-activation account of persuasion across different studies, and across different variables relevant to persuasion. On the other hand, both of these simulations had been conducted with the same connectionist network. In Simulation 3, I held the to-be-simulated phenomena constant, but introduced changes to the network.

Simulation 3

To demonstrate the network's robustness against various changes to the simulation parameters, I constructed an extended version of the network (see Figure 5).

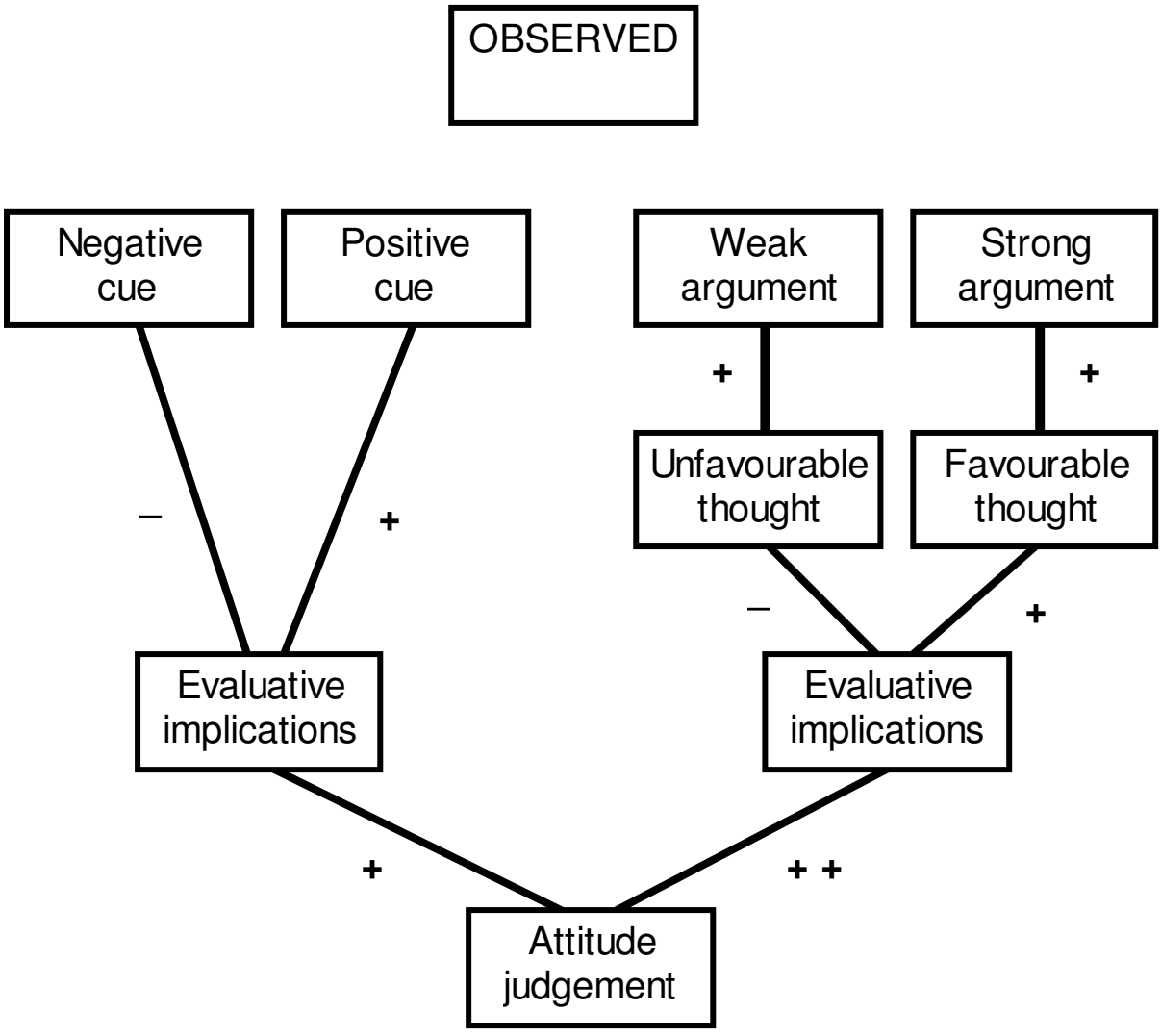

Figure 5. The parallel-constraint-satisfaction network used in Simulation 3. Boxes represent concepts, lines represent connections between concepts, plus- (minus-) signs depict mutual excitation (inhibition) of concepts.

Specifically, I added one unit to each of the network's paths (labeled

"evaluative implications" in Figure 5). These units may be interpreted as representing more specific evaluations than the global attitude judgement unit, namely the evaluative implications of the cue, and of the message, respectively (corresponding approximately to the variables "perceived communicator credibility" and "perceived message favourability" assessed by Chaiken and Maheswaran, 1994, as well as Maheswaran and Chaiken, 1991). The introduction of these units increased the length of each path, and also increased the total number of both units and connections participating in the network. Further, it eliminated any direct inputs of cue- or cognitive-response-units into the attitude unit. Note, however, that the introduction of these additional units was relatively neutral with respect to the assumed processing mechanism, as there were still fewer intermediate units involved in the spreading of activation from cue units to the attitude unit (one intermediate unit) than from argument units to the attitude unit (two intermediate units). Thus, if previous simulation results had been due to the assumed mechanism, the extended network should produce qualitatively comparable outputs. If previous results had been due to some other factor, previous results were less likely to replicate.

As before, a preliminary test with connection weights of equal magnitude but different sign (excitatory connections: .05; inhibitory connections: -.05) revealed that network outputs were dominated by the activated cue. The effect had been compensated for in the previous network by giving connections between cue units and the attitude unit a less strong weight. In the extended network, I instead increased the weight between attitude unit and the unit reflecting evaluative implications of arguments (using 1.5 times the default value, or .075 - depicted by a double plus-sign in

Figure 5). In effect, compared to the previous network, a smaller number of connection weights was changed from defaults, the change affected a different connection and, finally, was into the opposite direction (i.e., an increase, instead of the previously used decrease). Despite these deliberate differences at the implementational level, the two networks were *conceptually* comparable in that the strength of connection weights gave activation spreading from arguments an advantage over activation spreading from cues. Therefore, if previous simulation results had actually been due to the assumed processing mechanism, they should replicate. If they had been due to some other detail of the implementation, replication was less likely.

Simulation approach.

I adopted the same mapping of variables as in Simulation 2. Specifically, low (high) source credibility was considered to be a negative (positive) cue, ambiguous messages were simulated by externally activating both argument units, and processing under conditions of low (high) motivation was modelled by using network outputs from an early (late) stage. As the network was larger than the one used in Simulations 1 and 2, I redetermined the particular update cycles considered to represent low- *vs.* high-motivation processing. All simulations were run for 1000 update cycles. As no changes to unit activation levels occurred in or after update cycle 500, I adopted, for all simulated conditions, unit activations from cycle 500 as the network's high-motivation output. For some conditions, a reversal in sign was expected for the attitude judgement unit's level of activation. Preliminary tests revealed that this unit reached its greatest asymptotic activation with the

initial sign around update cycle 60. Therefore, I adopted activation levels from cycle 60 as the network's low-effort output for all simulated conditions.

Results

Simulation of attitude judgements. Like in the previous simulation, simulated attitude scores were qualitatively in line with expectations. Specifically, scores from early stages of processing reflected cue valence (see Table 11, left panel). Scores from late stages of processing reflected argument strength when conditions featuring unambiguous messages were simulated, but reflected cue valence when simulating conditions with ambiguous messages (see Table 11, right panel). Overall, scores were more extreme than in Simulation 2.

Table 11. Activation level of the attitude judgement unit as a function of cue valence and three levels of argument quality, at two stages of processing (Simulation 3).

		Processing stage					
		early			**late**		
		Argument(s)			Argument(s)		
		weak	ambig.	strong	weak	ambig.	strong
Cue	negative	−.56	−.47	−.16	−.62	−.61	.40
	positive	.16	.47	.56	−.40	.61	.62

Comparison of attitude scores with human data. Collapsing over cue, Chaiken and Maheswaran (1994) reported means of six average attitude judgements involving all twelve conditions of the design. Simulation scores correlated positively with these data from human participants, $r(4) = .99$, $p < .001$ (for comparison, previous network: $r = .98$; qualitative prediction pattern: $r = .94$). Further, the authors reported means of four average attitude judgements for each of two renditions of the ambiguous message. For each message rendition, simulation scores correlated strongly with human data, both rs(2) = .98, both ps < .03 (for comparison, previous network: rs were .87 and .96; qualitative pattern: rs are .99 and .96). Thus, if anything, then the extended network had a somewhat better fit with human attitude data than the previously used network.

Simulation of thought valence. As with the previously used network, both the unfavourable-thought unit and the favourable-thought unit showed positive activation at an early stage of processing when both argument units were activated externally. Again, the magnitude of activation was greater for the cognitive-response unit that was evaluatively congruent (as opposed to incongruent) with the cue unit activated simultaneously (early-stage activation scores of congruent and incongruent units were .46 and .22, respectively). In the previous network, the asymmetry of activation levels had become greater with extended processing due to the level of activation becoming more extreme for the congruent unit, but less extreme for the incongruent unit. In the present network, the asymmetry also increased with further processing. The increase was essentially driven by the activation level of the incongruent unit changing the sign (late-stage activation scores of congruent and incongruent units were .55 and -.13, respectively). As before,

the difference score (or: thought favourability index - see Table 12) was of greater magnitude in late than in early stages.

Table 12. Index score of thought valence as a function of cue valence and three levels of argument quality, at two stages of processing (Simulation 3).

		Processing stage					
		early			**late**		
		Argument(s)			Argument(s)		
		weak	ambig.	strong	weak	ambig.	strong
Cue	negative	−.87	−.24	.56	−1.06	−.68	1.04
	positive	−.56	.24	.87	−1.04	.68	1.06

Comparison of thought valence with human data. Difference scores from the simulation correlated positively with human data, $r(10) = .94$, $p < .001$ (for comparison, previous network: $r = .97$; qualitative pattern: $r = .89$).

Further results. Finally, I computed the correlation between attitude scores and scores on the valenced thought index within the simulation data, separately for early versus late stages of processing. As expected, a greater positive correlation was found for late than for early stages (late stage: $r(4) = .95$, $p < .004$; early stage: $r(4) = .64$, *ns*). The difference was marginally significant, $z = 1.63$, $p < .06$, one-tailed, for the test of difference between independent *r*s from the same sample (for comparison, the same result was found with the previous network; the correlation of the qualitative prediction

patterns under low task importance cannot be computed because of one of the variables involved being a constant).

Discussion

Simulation 3 introduced an extended version of the network used in Simulation 2. The extended network differed from the previously used network with respect to various simulation parameters. However, the extended network relied on the same processing mechanism, namely a greater number of units mediating the spreading of activation between attitude and argument units than between attitude and cue units. Several minor differences notwithstanding, simulation results were basically the same as in Simulation 2. Specifically, in both Simulation 2 and Simulation 3, simulated attitude scores had the expected sign in all conditions; both simulated attitudes and simulated thought valence correlated strongly with human data; finally, within the simulation data, these variables were versus were not correlated with each other as a function of processing time allowed, in the expected way. These findings are better explained by the common factor of the networks used, namely the underlying processing mechanism, than by details of the respective implementation.

General Discussion

I proposed the spreading of activation in networks of associated concepts as the cognitive operation whereby people select from, and integrate, information that is simultaneously available from a multitude of sources in persuasion settings. To demonstrate the viability of the approach,

I first derived a particular network of concepts from a discussion of current theories in the domain. Specifically, the network's structure reflected a basic assumption that seemed to be shared by other persuasion theories. In this network, I simulated persuasion phenomena by means of spreading activation.

I first simulated the classical finding of Petty, Cacioppo, and Goldman (1981) whereby low issue involvement leads to attitudes reflecting the valence of a context cue, but high involvement leads to attitudes reflecting argument quality (Simulation 1). In an extended replication (Simulation 2), I tested the same network against the additional conditions that Chaiken and Maheswaran (1994) introduced to demonstrate biased systematic processing. In addition to simulated attitude judgements, I recorded network outputs for cognitive responses. Finally, I replicated the second simulation with a network that was conceptually similar to the one previously used, but differed with respect to various implementational details (Simulation 3).

For both attitudes and thoughts, I found an excellent fit of simulation data and data from human participants (as reported in the literature), both qualitatively and quantitatively. Further, within the simulation data, the correlation between attitudes and thought valence was found to be greater at late (as opposed to: early) stages of processing. Thus, the network model did not only produce appropriate scores, but also correctly reflected the relation between variables. As the last simulation showed, these results did not depend on a specific set of simulation parameters. Taken together, results strongly support the viability of a spreading-activation approach to persuasion.

On the other hand, an excellent quantitative fit with human data was

also found for the qualitative prediction patterns derived from other persuasion theories. Whereas the connectionist model's output compared to human data no worse than the qualitative patterns, it also did not fit better. What, exactly, does the spreading-activation account contribute to our understanding of persuasion?

Together, the simulations showed that a range of basic findings from persuasion can be parsimoniously explained from only few, well-specified theoretical assumptions:

Assumptions shared by all models. The network model produced the target phenomena from a single cognitive mechanism. The mechanism comprises two basic assumptions shared with other models, namely a) that persuasion by arguments is mediated by cognitive responses, whereas persuasion by cues is not (or: less so), and b) that low-effort processing versus high-effort processing differ (at least) with respect to the amount of processing that occurs overall.

Cognitive operations. Both the the unimodel and the spreading-activation account postulate a single, specific cognitive operation as underlying persuasion throughout. From the two cognitive operations, are there reasons to prefer one over the other? Kruglanski and Thompson (1999a) showed that the impact of both cues and arguments on attitudes may be modelled as syllogistic reasoning. For the spreading of activation, the present simulations made the same point. In addition, and orthogonally to the question how cues and arguments exert their impact, the present simulations showed that spreading activation may account for the outcomes of both low-effort and high-effort processing strategies. It is not clear if a similar demonstration can be provided for syllogistic reasoning. Specifically,

Kruglanski and Thompson (1999a) refer to "memory search and activation processes" (i.e., the search for relevant major premises in memory, and their activation beyond some functional threshold - p. 90) or "cognitive work" that may become necessary if evidence is unclear or complex. These processes are not specified in much detail. Therefore, it remains to be shown if they can be described exclusively in terms of syllogistic reasoning (if not, the unimodel would no longer be a single-process model). More importantly, however, the spreading-activation model does not require an active search process to begin with. Instead, activation is allowed to spread indiscriminately through the network. If no direct link exists between two concepts (as was the case for arguments and attitudes), indiscriminate spreading will result in activation spreading between them more indirectly, via other concepts that happen to be associated with both of them. Thereby, although no active search mechanism is involved, mediating concepts (or: cognitive responses, i.e. the equivalent of the unimodel's major premises) will become activated nevertheless. - In sum, both syllogistic reasoning and spreading activation seem to be viable cognitive operations that may underlie persuasion. Spreading activation appears the more parsimonious one as it does not require the assumption of dedicated memory search processes.

Cognitive processes. In line with dual-process models, the present approach relies on a systematic difference in the number of concepts mediating the attitudinal effects of cues versus arguments. The question whether this difference constitutes a quantitative or a qualitative one is subject to debate in the literature. More importantly, however, the difference proved sufficient to allow for the simulation of basic findings in persuasion. Adopting spreading activation as the network's single cognitive operation, the

effects of both low-effort and high-effort processing (as well as biased systematic processing) on both attitudes and thought valence (as well as their association) could be modelled. Thus, the spreading-activation approach is specified in sufficient precision to allow for a computer simulation of persuasion phenomena.

A within-subjects mechanism. Finally, note that although the network was used to simulate between-subjects data, the network model's cognitive mechanism basically describes a within-subjects process. Thus, the model allows for the prediction of attitudes in line with cues at an early stage of processing, even if people process under conditions of high motivation and capacity, and will ultimately arrive at an opposite judgement. In the next chapter, I will present steps towards the development of a technique that allows to show the postulated within-subjects effect directly with human participants.

Chapter 3: An instrument for observing the construction of a social judgement

In this chapter, I describe first steps towards the development of an instrument for the repeated assessment of judgements at predetermined stages. I first give a general introduction to the technique. Then, two laboratory studies using the instrument are presented. The studies tested various technical parameters, but also dealt with possible carry-over effects, or "contamination" of the second judgement by the first one.

Introduction

Various theoretical models in social cognition describe social judgement as the result of two distinct cognitive processes. One of these processes is usually assumed to be efficient, effortless, and relatively automatic, whereas the other process is typically assumed to require more cognitive effort, and to be more consciously controlled (E. R. Smith, 1994; Smith & DeCoster, 2000). More often than not, these dual-process models hold that both of the processes may occur together during the construction of a social judgement. In fact, in their discussion of ten major dual-process approaches, Smith and DeCoster identified only two (important, although not exactly recent) models assuming the opposite, namely the mutual exclusiveness of processes (see Smith & Decoster, 2000, p. 125).

Whereas dual-process models typically postulate within-subjects processes, laboratory demonstrations of these processes and their effects

have usually relied on between-subjects designs. That is, the effects of each process have been shown in different participants (as opposed to: in the same participant). For a broad range of examples of research using between-subjects designs in the domain of persuasion see, for instance, the monograph by Petty and Caccioppo (1986) and the textbook by Eagly and Chaiken (1993). Below, I will report the development and test of an instrument that allows for the demonstration of the two processes *within-subjects*. Before doing so, however, I should discuss why a within-subjects demonstration might be desirable to begin with. To do that, I first describe a study from the dual-process literature where the traditional between-subjects approach was taken. Then, I outline how the assumed processes might be shown within-subjects, and further, what the specific advantages of the within-subjects approach are.

Advantages over the traditional approach

Knowles and Condon (1999) have recently drawn on the two-stage Spinozan belief model (Gilbert, 1991) whereby information is automatically accepted in an initial, effortless comprehension stage, and cannot be rejected except in a subsequent, more effortful reconsideration stage. In their test of the proposed within-subjects mechanism, Knowles and Condon (Studies 2 and 3) adopted a between-subjects cognitive-load manipulation. Cognitive load was expected to interfere selectively with the ability to reconsider information, but not with its automatic acceptance. Thus, participants in no-load conditions should be able to move on from automatic acceptance to reconsideration. Participants in cognitive-load conditions, in contrast, should

not be able to go beyond the stage of automatic acceptance. Consistent with this reasoning, Knowles and Condon predicted greater acceptance for cognitive-load conditions than for no-load conditions. The prediction was supported in both studies, a finding in line with the proposed dual-stage mechanism.

As an alternative to Knowles' and Condon's between-subjects approach, consider a fictitious within-subjects study that does not adopt a cognitive-load manipulation. Instead, in this fictitious study, judgements are assessed repeatedly: first, when participants conclude the automatic-acceptance stage, and next, when they conclude the reconsideration stage. Obviously, there would be quite a few difficulties involved in actually running such a study. I will discuss these difficulties (as well as ways to overcome them) in more detail below. Assuming for a moment that the study can be run as described, the within-subjects demonstration of dual stages has at least three advantages over a between-subjects approach: it is *more economical*, provides a *more conclusive* test of the underlying process assumptions, and is *less vulnerable* to alternative explanations.

Most obviously, as the between-subjects cognitive-load manipulation is not used, the described within-subjects study can be run with half the sample size, rendering it *more economical* than a between-subjects approach. Alternatively, if participant numbers are not a concern, the same sample size may be used in order to increase the statistical power of hypothesis tests.

Next, compared to a cognitive-load manipulation, repeated assessment appears to be the *more conclusive* test of the two-stage model. We may plausibly infer that Knowles' and Condon's no-load participants had

to go through an initial automatic-acceptance stage before they entered the reconsideration stage. However, technically speaking, this has not been *shown*, as each participant's judgement was assessed at one point in time only. With repeated assessment, in contrast, participants' judgements from two points in time are available. Thus, changes in acceptance over time, if they occur, can actually be *observed* (as opposed to: inferred).

Finally, the introduction of a cognitive-load manipulation into Knowles' and Condon's (1999) Studies 2 and 3 also introduced a possible alternative explanation. According to the alternative explanation, the presence of cognitive load triggers acceptance, whereas the absence of cognitive load does not. Note that the alternative explanation accounts for the findings more parsimoniously than the proposed Spinozan mechanism, as it omits the postulation of two stages. To be sure, Knowles and Condon did provide further evidence for the process assumptions in their Study 1. When taken together, their three studies do exclude the alternative explanation. However, by not introducing a between-subjects manipulation in the first place, the within-subjects approach avoids alternative explanations revolving around that manipulation to begin with. Alternative explanations revolving around the repeated assessment, in contrast, appear somewhat difficult to construct: An explanation such that first assessment triggers acceptance, whereas repeated assessment does not, would not exactly be *alternative* to the two-stage mechanism, but would actually be in line with it. Thus, the within-subjects approach appears *less vulnerable* to alternative explanations.

Taken together, a within-subjects approach to demonstrating dual stages (or dual processes) appears desirable. The basic idea of the approach is quite straightforward: judgements are to be requested twice -

first, when they reflect the impact of one of the processes more strongly, and next, when they reflect the impact of the other process more strongly. I will discuss three questions that arise when trying to translate the basic idea into an experimental paradigm. The first question concerns the timing of the two requests - when, exactly, should they be delivered? The second question concerns the timing of participants' responses - how can participants be encouraged to respond immediately to the requests? The third question concerns the effects of requesting responses repeatedly – does the method unduly influence the course of cognitive processing?

The timing of judgement requests: Accounting for differences between participants, and between items

A straightforward approach to the experimenter-controlled timing of responding is the adoption of static deadlines, for example, by Wilson, Lindsey, and Schooler (2000, p. 114): "... half of the participants were given only 3 s to respond ... , whereas the other half were given 30 s". This example describes a between-subjects manipulation, but could analogously be used in a within-subjects design. I decided not to do so as the method appears rather brittle. In particular a static lower-bound value is likely to require frequent adaptations when changing stimulus materials, or may not pose comparable demands on different participants. Instead, I considered a more dynamic approach to determining the timing of requests preferable.

Specifically, I adopted two theoretically defined values. Dual-process models assume one process to be relatively effortless and automatic, and the other process to require more cognitive effort. If so, then the effects of the

effortless process on judgements should be most pronounced when the stimulus material has just been comprehended, but has not been processed much further. I assumed that this point in time would correspond approximately to the time required for *reading* a stimulus item. The effects of the more effortful process, in contrast, should be most pronounced when participants have expended all the cognitive effort they are willing to expend in the experimental situation. I assumed that this point in time would correspond approximately to the time participants take, under self-paced conditions, before *reporting an evaluative judgement* of an item. Thus, the *first request* should occur between the time required for reading an item, and the time required for evaluating an item (in a self-paced task). The *second request*, in contrast, should occur at or after the point in time where the reporting of a judgement is to be expected anyway. - Both reading-time and self-paced judgement-time are, for a given participant and stimulus item, not known in advance. It should however be possible to *estimate* them from a participant's time requirements when reading and when judging other, similar items at their own pace.

Independent of individual differences, participants' response time requirements are affected by arbitrary differences between items, for instance by an item's wording (Knowles & Condon, 1999, Studies 2 and 3; and see my Studies 2 and 3 below), or by its length (number of words). In a study on the schematic effects of attitudes, Judd and Kulik (1980) found a significant, unique contribution of item length to the prediction of response times in an evaluative task (see their Table 1, p. 573). For the present research, it appeared useful to take item length into account when determining the timing of judgement requests (see the Methods sections of

Studies 2 and 3 below for details).

The timing of judgements: Synchronising request and response

Presenting a response request at a certain point in time may be followed by a response in less or more close temporal proximity to the request. Draine and Greenwald (1998) described various response criteria that participants may adopt in studies on priming effects. Accordingly, participants may trade-off speed for accuracy. That is, they may respond more slowly than they could, in order to avoid errors. Or, vice versa, they may tolerate higher error rates than necessary, in order to respond more quickly. The criteria adopted may differ between participants, or may differ within the same participant at various points in time. The problem arising in priming studies is that adoption of different criteria leads to a dilution of priming effects across measures of response speed and measures of response accuracy, rendering it more difficult to observe effects on either measure.

A related problem was likely to occur in a within-subjects demonstration of two processes. Specifically, participants may delay a response to the first request, in order to avoid answers that are potentially not in line with the results of subsequent processing. Thereby, judgemental effects of the effortless, relatively automatic process would become distributed across response speed and measures of the manifest judgement, rendering it difficult to find effects on either dimension.

To control for trading-off, Draine and Greenwald (1998) adopted a response window technique whereby a response prompt was presented at a

predetermined time, signalling that a response should now be made. Importantly, the prompt had a limited duration, thereby also introducing a response deadline. In effect, the use of response windows constrained participants' choice of when to respond, and returned control to the experimenters. I adopted the response window technique for the present studies because it appeared useful for preventing the possible delaying of responses to first prompts by participants.

Repeated assessment: Testing for a processing bias possibly introduced by the method

Further, I was interested in the instrument's effects on processing. Specifically, it is conceivable that the overt reporting of a preliminary judgement at an early stage (that is, before a judgement would be reported anyway) may change the course of subsequent, effortful processing. If so, then the instrument would undesiredly introduce a processing bias. To test for this possibility, it is necessary to compare high-effort judgements that *were* versus *were not* preceded by the overt reporting of a low-effort judgement. This might be accomplished in a study where half of participants are presented with response prompts both before and after the time they would choose to report a judgement anyway, whereas the other participants are presented only with the latter of these prompts. If judgements in response to the only (single-prompt condition) versus the second (repeated-prompts condition) response signal were found to differ, a processing bias introduced by the instrument would have been shown. If they were found not to differ, no evidence for a bias would have been found. Note, however, that in the

experimental design just described, the demonstration of the instrument not introducing a processing bias would rely on a null effect.

The prediction of a null effect can be avoided, and a more conclusive test for the presence or absence of method-induced bias can be provided by extending the described design so that it allows for the test of an interaction. Consider a design where, for half of participants, a response signal is given before the point in time they would choose for responding in a self-paced task. For the other participants, the signal is given only afterwards. The design allows for a between-subjects test of differences between judgements as a function of time of prompt onset. Actually, we should expect to find differences, because when responding, processing time had been ample for participants in one condition, but not in the other. Now assume that an additional response signal was presented, for both conditions at the same, very late, point in time. For judgements elicited by the additional prompt, processing time would have been ample for all participants, and no difference in judgements should be found between conditions. In this design, the absence of method-induced bias would be indicated by a significant interaction such that first judgements differ as a function of time of prompt onset (indicating that the instrument was effective), whereas second judgements do not differ (indicating that overt reports of preliminary judgements do not trigger changes in subsequent judgements). - The studies to be reported below featured the experimental manipulations required to test for this interaction.

Selection and pretest of materials for Study 2

Study 2 was to adopt a persuasion paradigm. Specifically, I planned to present participants with statements of a known valence, in combination with heuristic cues of a known valence. With respect to statements, I tried to find both an agreeable and a disagreeable statement on various topics, in order to avoid confounding of statement valence and statement content (or: topic). As heuristic cues, I tried to identify brief person descriptions that would trigger the expectation that a statement from that person was likely to be agreed versus disagreed with.

Opinion statements

I prepared a pool of 24 opinion statements that covered a broad range of general-interest topics and expressed, in one sentence, an opinion on the topic (e.g., "Genetically manipulated food is an invention that should be saluted instead of despised."). The statements were adapted from actual opinions expressed on the BBC website, or were drawn from published scales. For each statement, I created a complementary statement that expressed the opposite view (e.g., "Genetically manipulated food is an invention that should be despised instead of saluted."). Then, two renditions of a statement list were created that contained a different member from each of the 24 complementary pairs of statements. Further, renditions were compiled such that they contained approximately equal numbers of statements appearing likely to be agreed versus disagreed with.

Person descriptions

Next, a list of 32 person descriptions was prepared. These descriptions comprised a first name as well as a label of one or two words. Approximately equal numbers of female and male first names were used. Most of the labels represented a profession; some of the labels pointed to a person's age or geographical location. Examples are: "Ben (farmer)", "Laura (57 yrs)", "Hazel (Aberdeen, UK)".

Pretest procedure

The lists of opinion statements and person descriptions were then submitted to a pretest. Twenty-six second-year and third-year psychology students participated in partial fulfillment of course requirements. Thirteen participants each rated the different renditions of the statement list. All participants subsequently rated the list of person descriptions. As the materials should later be used in studies adopting time pressure manipulations, pretest instructions for both the opinion-statement rating task and the person-description rating task explicitly encouraged participants to respond quickly, and to rely on their first impressions.

Statements. Instructions pointed out to participants that they were to indicate their level of agreement or disagreement with opinion statements, or one-sentence summaries of opinions that had been raised on the Internet, in newspapers, or other sources. In order to encourage quick responding, instructions emphasised that the research interest was on "spontaneous reactions" towards a statement. For the same reason, participants were informed that they were to rate forty opinions overall, and that the rating of

opinions would be followed by other tasks. The 24 statements, as well as 16 similar filler statements that did not differ between list renditions, were then presented on the computer screen, one at a time, in a different random order for each participant. Participants responded on seven-point scales with the title "I agree ..." and the endpoints labeled "not at all" (1) and "very much" (7). Reporting level of agreement by pressing one of the keys from 1 to 7 triggered a 0.5 second blank-screen interval, followed by presentation of the next statement, or, after the last statement, by presentation of the instructions for the person-description rating task.

Person descriptions. Instructions stated that this part of the study was interested in what can be inferred about forthcoming opinion statements from extremely short person descriptions. The possibility of such inferences was illustrated using the person descriptions "Joseph (7 yrs)" and "Susan (professor)", and referring to the different expectations about the *level of sophistication* of a forthcoming statement that these descriptions may raise. Then, participants were informed that about 100 judgements were to be made, and that they were strongly encouraged to proceed quickly once they had started the task. The list of person descriptions was randomly reordered for each participant. Then, for each person description, three screen displays were constructed by completing the template "Assume that [person description] had made a statement on a topic of general interest. Generally speaking, how much would you expect that statement to be ... [attribute]". The three attributes used were "reasonable", "in touch with reality", and "in line with your own views". The first and second attribute served as distractors. The third attribute ("in line with your own views") was considered most decision-relevant for participants in subsequent studies who would,

under time-pressure, have to decide whether to agree or disagree. The completed templates were then presented on the computer screen, one at a time. Participants responded on seven-point scales with the endpoints labeled "not at all" (1) and "very much" (7). Reporting level of expectation by pressing one of the keys from 1 to 7 triggered a 0.5 second blank-screen interval, followed by presentation of the next person/attribute combination, or, after the last combination, by debriefing instructions.

Pretest results

Statements. For each of 24 statement pairs, I tested agreement scores of the constituent statements for difference from the scale midpoint (4), using one-sample t-tests. A pair was retained if scores for one of the statements were significantly *below* the scale midpoint, indicating disagreement, whereas scores for the other statement were significantly *above* the scale midpoint, indicating agreement. Seven pairs of statements were found to unequivocally meet this criterion (see Appendix B), with the respective two *t*-values each being both opposite in sign and more extreme than -3 and 3, *df*s = 12, *p*s < .01.

Person descriptions. For each person description, I tested scores on the target attribute ("in line with your own views") for difference from the scale midpoint (4), using one-sample t-tests. Three person descriptions were found to be unequivocally *above* the scale midpoint, indicating the expectation of agreement: "Sharon (student)", and "Mark (student)", with *t*s > 4.1, *df*s = 25, *p*s < .001, and further "Nadine (21 yrs)", $t(25) = 3.71$, $p < .002$. Three person descriptions were found to be unequivocally *below* the scale midpoint,

indicating the expectation of disagreement: "Wendy (13 yrs)", "Paul (15 yrs)", and "Mike (bouncer)", *t*s < -4, *df*s = 25, *p*s < .001. Other *t*s were in the range from -2.6 to 2.3, *df*s = 25, *p*s > .01.

Study 2

Overview

Major research questions

In Study 2, participants were presented with items comprising a heuristic cue (a person description) and a statement. The valence of both cue and statement was known from pretest results, and was orthogonally varied. In line with theories of persuasive communication, I expected participants' first judgements to reflect cue valence, but expected their second judgements to reflect statement valence. Importantly, this should only be the case if first judgements were requested before (as opposed to: after) the point in time where participants would report a judgement spontaneously. Finally, the possibility of a method-induced processing bias was investigated.

An additional factor

Similar to the present studies, Knowles and Condon (1999, Studies 2 and 3) adopted one-sentence statements as stimulus material, as well as a dichotomous response format ("yes" versus "no"). The authors found (Study 2) a replicable (Study 3) question type bias, such that self-descriptive statements were more often endorsed when they were phrased as assertions

(e.g., "I am extraverted") than when they were phrased as negations (e.g., "I am not introverted"). For the online algorithm adopted in the present studies, there is no difference between assertions and negations except for number of words. It appeared useful to investigate whether the repeated-measures instrument would be affected by statement type (assertion versus negation). Therefore, Study 2 featured both assertions and negations.

An additional measure

Finally, a new measure was to be explored in Study 2. As described in the introduction, in speeded response tasks, participants may trade off speed for accuracy - that is, they may respond somewhat more slowly than they could, in order to avoid response errors. The explicit purpose of using response windows is to reduce the overall amount of trading-off, by constraining participants' choice of when to respond. Note, however, that the technique has initially been developed in the context of research using one-word stimuli, and adopting window durations in the magnitude of, for instance, 133 milliseconds (e.g., Draine & Greenwald, 1998, Study 1). The present line of research, in contrast, adopted one-sentence statements as stimulus material. As a consequence, more lenient window durations were employed. Therefore, it was not unlikely that trading-off might still occur *within* response windows, distributing, as Draine and Greenwald pointed out, the effects of experimental manipulations across two variables (here: speed of responding, and manifest judgement), and rendering it less likely to find differences on any one of them. As these authors further pointed out, combining distinct measures of speed and of accuracy into a single one

would benefit from knowing participants' speed-accuracy trade-off function. This function is, of course, typically unknown. On the other hand, knowledge of the *actual* speed-accuracy trade-off function may not be necessary for effectively improving measurement in the present context. Instead, I assumed that it would be possible to find a *useful approximation* of the function. Specifically, I adopted a particular multiplicative combination of response valence and response latency (see the Methods section for details) as a single index reflecting both of these dimensions simultaneously. Should the effects of experimental manipulations in fact become diluted across speed and valence of responding, then I expected the measure simultaneously reflecting both speed and valence to be affected more strongly than measures of any one of the dimensions alone.

With the additional factor and additional measure, the design of Study 2 was quite complex. I will therefore summarise my hypotheses only at the end of the Method section. Although it is not usual to present hypotheses there, it will be more easy to understand them when familiar with design and variables.

Method

Participants

One-hundred and twenty-seven first-year psychology students (108 female, 19 male) participated in one of three mass-testing sessions as part of an introductory course on experimentation and research methodology.

Materials

Target items. Based on the pretest results described above, four person descriptions were selected to serve as heuristic cues. Specifically, "Sharon, student" and "Mark, student" were employed as cues suggesting that statements would likely be in line with participants' own views, whereas "Wendy, 13 yrs" and "Paul, 12 yrs" were used as cues triggering the opposite expectation. - Seven pairs of statements of opposite valence had been identified in the pretest. Two of them were considered inappropriate for the present study because they dealt with children, and were thus related in content to the cues relying on young age ("Wendy", and "Paul"). From the remaining five pairs, four were selected for inclusion in the study. With the help of a native English speaker, I rephrased these statements such that statement *valence* (agreeable versus not agreeable) and statement *wording* (assertion versus negation) were unconfounded. See Table 13 for the set of statements actually used in the study. - From person descriptions and opinion statements, four *target items* were constructed by appending to each person description the appropriate statement (depending on experimental condition; see below).

Table 13. Target opinion statements (Study 2)

Positive valence assertions

"Writing and spreading computer viruses is wrong even if no harm is intended."

"People who live in areas at risk for natural disasters should be helped by the government when a disaster strikes."

Negative valence assertions

"Giving directly to beggars, rather than to charities, is the best way to help the needy."

"If a friend came into financial difficulty you should refuse to lend them some money."

Positive valence negations

"Giving directly to beggars, rather than to charities, is not the best way to help the needy."

"If a friend came into financial difficulty you should not refuse to lend them some money."

Negative valence negations

"Writing and spreading computer viruses is not wrong as long as no harm is intended."

"People who live in areas at risk for natural disasters should not be helped by the government when a disaster strikes."

Note. Statements depict the assertions as well as the negations of positive as well as negative valence used in Study 2.

Practice and filler items. Two additional lists of 34 person descriptions and of 34 opinion statements, respectively, were compiled from pretest materials. For each participant, the list of person descriptions was reordered

randomly. Then, 34 practice and filler items were constructed by appending to each person description the statement of the same list position.

Item formatting and display. For both practice and target items, person descriptions and statements were separated by a colon. Statements were surrounded first by blanks, and further by quotation marks. Items were displayed within a white box on an otherwise light-grey computer screen, using horizontally centred alignment. Person description and colon appeared in the first line of the box. The statement appeared in the second and, where necessary, third line.

Presentation order. From the 34 practice and filler items, ten items were used for the assessment of reading speed, and another ten items were used to assess evaluation speed. The items used in the reading-speed task were also employed for practicing responding to response prompts. The remaining 14 items served as fillers in the repeated-judgement task. Specifically, the four target items were appended to this list of 14 filler items. For each participant, the last eight (out of now 18) items in this list were reordered randomly. Thus, in effect, the repeated-judgement task started with ten filler items, followed by four filler items as well as four target items in a random order.

Procedure and instructions

Participants were informed that the study was about the perception of others' views and opinions under conditions where information is very scarce. They further learned that they would be presented with brief, slogan-like opinion statements collected from the Internet, from newspapers, and from

other sources. By pressing a key, participants started the four-step experimental procedure. At the beginning of each step, instructions briefly described the respective task. Further, they encouraged participants to keep fingers on the respective response keys between responses.

In *Step 1*, participants read, at their own pace, a filler item presented inside a white box on an otherwise light-grey screen, and pressed a dedicated key afterwards. The keypress cleared the white box. After a one-second interval, the next filler item appeared inside the box, or, after 10 filler items, instructions for the next step were shown. In Step 1, an index of a participant's reading speed in self-paced tasks was determined (see below for details).

In *Step 2*, participants reported, at their own pace, their agreement or disagreement with a filler item presented inside the white box. Specifically, they pressed one of two dedicated keys in order to indicate that they either "tend to agree" or "tend to disagree" with the statement. Pressing one of these keys cleared the white box. After a one-second interval, the next filler item appeared inside the box, or, after 10 filler items, instructions for the next step were shown. In Step 2, an index of a participant's evaluation speed in self-paced tasks was determined (see below for details).

Step 3 introduced response windows. Specifically, participants again reported their agreement or disagreement with filler items by pressing one of two dedicated keys. This time, however, they were asked to respond only while a blue frame was visible. They were told that correctly having timed a response would be indicated by the blue frame turning green. The first practice item was then shown inside the white box. After some time (see below for timing parameters), a blue frame occurred for two seconds,

surrounding the white box. If a participant responded while the frame was visible, the frame changed its colour from blue to green for the remainder of its two-second duration. If a response was made too early or too late, a brief message explained the error, and the trial was repeated with the same item. After all of ten practice items had been responded to in time, the next step was started.

In *Step 4*, participants again reported their agreement or disagreement with items, with the timing of responses being controlled by the presentation of blue frames. This time, however, *two* response prompts were shown for each item. Depending on experimental condition, the first of these was shown either *very early* or *not very early* (see below for timing parameters). Constant across conditions, the second prompt was shown very late. The dependent variables were assessed in Step 4. Importantly, from the 18 items used in this step, the first ten were practice items. The four target items appeared only afterwards, at random positions among the last eight.

Next, participants filled out three personality scales. This served mainly as a distractor task and was followed by the assessment of manipulation checks. Specifically, participants answered questions created from the template: "Assume you were going to hear a statement from a person introduced as [person description]. Generally speaking, how much would you expect that statement to be in line with your own views?". Among others, the four target person-descriptions used in this study were filled in. Participants reported their level of expectation on seven-point scales with the endpoints labeled "not at all" (1) and "very much" (7). Finally, participants' gender, age, native-speaker status, and self-ascribed proficiency of English were assessed. Participants were then thanked and debriefed.

Timing parameters

Speed indices. For each of ten responses collected in the reading-speed task, a word-based speed index was computed by dividing the respective *response time* by the item's *number of words*. From these ten indices, the two smallest, as well as the two largest, were discarded because of potentially being outliers. The remaining six indices were averaged to compute a participant's overall reading speed index. - An equivalent procedure was followed to compute a participant's overall evaluation speed index.

Response window duration. Response prompts ("blue frames") had a constant duration of two seconds.

Onset of response prompts during training. In the training task, individual response latencies for each item were predicted by multiplying the evaluation speed index with the word number of the item. A response signal was given one second later. This timing appeared sufficiently lenient to facilitate learning of the task, in particular since training items were actually the same items that had been used for the assessment of reading speed.

Onset of response prompts in repeated assessment. As this was one of the first tests of the instrument, I adopted more lenient timing parameters than the instrument allowed for. To compute the onset of the *very early signal* for any given item, reading speed index and evaluation speed index were averaged, and the result was multiplied with the statement's number of words. Then, one second (50 percent of the signal's duration) was subtracted. Thereby, in effect, the centre of the response window coincided

with the average of estimated reading time and estimated evaluation time. This formula should guarantee that participants had sufficient time to read a statement and build a first impression, but not sufficient time to complete evaluative processes. To prevent unforeseen overlaps of first and second response window, the onset of the *not very early* response signal was not computed dynamically. Instead, it constantly started four seconds after the end of the first one. This figure, as well as the two-second duration of a response signal, was chosen based on the experiences collected in pretests. Finally, the *late* signal started constantly four seconds after the end of the *not very early* one. - Note that three prompt onsets were *computed* for each participant, although only two prompts were actually *shown*. Depending on experimental condition, these were either both the "very early" and the "late" one, or both the "not very early" and the "late" one. Thus, onset of the late prompt was independent of experimental condition.

Design

Study 2 featured a fully factorial, mixed design with three within-subjects factors and two between-subjects factors. The within-subjects factors were ordinal position of response (first vs. second response towards the same item), item cue valence (negative vs. positive), and item statement valence (negative vs. positive). The between-subjects factors were onset of the first out of two response prompts (very early vs. not very early) and item list used (assertions vs. negations).

Further, the assignment of cues to statements was controlled by four additional between-subjects factors. For instance, to construct a positive-

cue/positive-statement item, either "Sharon, student" or "Mark, student" could be used as the cue (factor 1), and either of two positive statements could be employed (factor 2). Independent of this, negative-cue/negative-statement items could be constructed using either of two negative cues (factor 3) and either of two negative statements (factor 4). These factors were adopted for counterbalancing purposes only. They will not further be discussed.

The major dependent variables were participants' repeated (first vs. second) agree/disagree-responses towards each of four (two [cue-valence: negative vs. positive] by two [statement valence: negative vs. positive]) target items, as well as the associated response latencies. Additional dependent variables were the scores on a measure of both speed and valence of responding (see below).

Scoring of dependent variables

Reversed response times. For each response, I computed the time remaining within the respective two-second response window after the decision had been made. For instance, if a response was made 500 ms after onset of the response signal, the score would be
2000 ms - 500 ms = 1500 ms. The theoretical range of these variables was 0 ms to 2000 ms, with greater values indicating quicker responses (i.e., more unused time remaining within the respective response window). Missed responses were assigned a score of zero, indicating that for the respective (non-)judgement, the available decision time was fully used up.

Dichotomous judgements. Participants' dichotomous judgements of agreement or disagreement were coded -1 ("I tend to disagree") and 1 ("I

tend to agree"). Missed responses were assigned a neutral score of zero.

Measure of both valence and speed. For each judgement, an index depicting both the response's valence and its speed was computed by multiplying the respective response score (-1, 0, or 1) with the value of the reversed response time (0 to 2000). These indices had theoretical ranges from -2000 to 2000. Scores close to the lower end of the scale (-2000) indicated quick disagreement, whereas scores close to the upper end of the scale (2000) indicated quick agreement. Responses that were made less quickly resulted in scores closer to the scale midpoint. Missed responses had a neutral score of zero.

Hypothesis

I expected a four-way interaction involving all of the within-subjects factors as well as the between-subjects factor "onset of the first prompt" such that first judgements should essentially reflect cue-valence, whereas second judgements should essentially reflect statement-valence, a pattern that should be most pronounced when the first prompt was presented very early (as opposed to: not very early). In contrast, I had no unequivocal predictions for the effects of judging assertions vs. negations.

Results

Missed responses

Participants missed twenty-four first responses (4.7%), and four second responses (0.8%). Three separate ANOVAs with number of missed

first, second, and first plus second responses, respectively, as the dependent variables, and both item list and onset of the first prompt as between-subjects factors were conducted. They revealed no systematic differences in numbers of missed responses, *F*s < 1 for item list as well as for the interaction of item list and onset of the first prompt; *F*s < 2.1, *p*s > .15 for onset of the first prompt. Thus, whereas participants missed more first than second prompts in terms of absolute frequencies, missed responses were nevertheless distributed evenly across the design's four between-subjects conditions.

Manipulation checks

I conducted a mixed-model ANOVA with the four manipulation checks as the dependent variables, onset of the first prompt and item list used as between-subjects factors, and cue valence as well as cue gender as within-subjects factors. As expected, the analysis revealed a main effect of cue valence such that positive cues (*M* = 4.52) triggered more strongly than negative cues (*M* = 3.27) the expectation that a not-yet-known statement from a person so described would be agreed with, $F(1,123) = 91.32$, $p < .001$, *MSE* = 2.16. Thus, in addition to differing from each other, the means of cue manipulation checks were found on expected sides of the response scale from 1 ("not at all") to 7 ("very much"). - Further, expectations of agreement tended to be smaller for male cues (*M* = 3.86) than for female cues (*M* = 3.94) overall, but the difference did not reach statistical significance, $F = 2.78$, $p < .10$, *MSE* = .31. No other effects occurred, *F*s < 2, *p*s > .16. Thus, by this criterion, the manipulation of cue valence was successful across the four between-subjects conditions.

Reversed response times

I conducted a mixed-model ANOVA with participants' eight response time scores (in milliseconds) each as dependent variables. The between-subjects factors were onset of first prompt (not very early vs. very early), and item list (negations vs. assertions). Cue valence (negative vs. positive), statement valence (negative vs. positive), and ordinal position of response (first vs. second) were the within-subjects factors.

The analysis revealed main effects of onset of first prompt, $F(1,123) = 10.32$, $p < .003$, $MSE = 253096$, and of ordinal position of response, $F(1,123) = 45.95$, $p < .001$, $MSE = 137916$. Importantly, these factors entered into an interaction with each other such that, for *first responses*, less time was left within a window when the response was requested very early ($M = 1184$ ms) than when it was requested not very early ($M = 1411$ ms), $p < .001$ for the pairwise comparison. The time available after *second responses*, in contrast, was not affected by the onset of the first prompt, $M_{very\ early} = 1467$ ms, $M_{not\ very\ early} = 1444$ ms, $p > .40$ for the pairwise comparison. The interaction just described was significant, $F(1,123) = 28.51$, $p < .001$, $MSE = 137916$. Other effects did not occur, Fs < 2.4, ps > .12.

According to these results, participants' first responses were slower when the first prompt occurred very early than when it occurred not very early, suggesting that the time pressure manipulation was successful. Importantly, as intended, the very early prompt selectively increased latencies of first responses, but did not exert an impact on latencies of

subsequent responses. Finally, it should be noted that response times were affected exclusively by manipulations of *process variables* (that is, onset of the first prompt, and number of response) but were insensitive to manipulations of *item content* (that is, cue valence, statement valence, and item list).

Dichotomous judgements of agreement

Participants' eight judgement scores each were submitted to the same mixed-model ANOVA using all experimental factors as described for response times. Different from response time results, the analysis revealed several main effects and interactions that were not relevant for a test of the research hypothesis. With 31 effects tested, and a large (in relation to the two-by-two between-subjects part of the design) sample size of 127 cases, this finding was not too surprising. In the interest of parsimony, I will focus exclusively on the highest-order effect relevant to the repeated assessment of judgements. See Appendix C1 for a comprehensive discussion of all ANOVA effects.

First of all, the predicted four-way interaction of the major experimental factors was not found, $F < 1$. The highest-order effect of interest for the present research was a marginally significant interaction of statement valence and onset of the first prompt. Accordingly, *negative-valence statements* were less disagreed with when the first prompt was presented very early ($M = -.71$) than when it was presented only later ($M = -.92$), $p < .001$ for the pairwise comparison. In contrast, *positive-valence statements* did not differ as a function of onset of first prompt ($M_{very\ early} = .57$,

$M_{not\ very\ early}$ = .55, p > .80 for the pairwise comparison). For the two-way interaction, $F(1,123) = 3.00$, $p < .09$, $MSE = .84$. Thus, with respect to negative-valence statements, there was some evidence that the prompt onset manipulation did affect judgements. However, the two-way interaction was not qualified by three-way interactions with either ordinal position of response ($F < 2.5$, $p > .11$), or cue valence ($F < 1$), or any higher-order interaction involving at least one of these factors (Fs < 1). Accordingly, whatever the effects of early prompt onset, they did not seem to affect first versus second responses differently, and did not seem to trigger an increased impact of cue valence on judgements.

To sum up, participants' dichotomous judgements were essentially determined by statement valence. Whereas some evidence for the effectiveness of the onset-of-first-prompt manipulation was found, ordinal position of response, as well as cue valence, did not have much of an impact on judgements. Thus, different from the response time data, the dichotomous judgement data actually provided little support for my hypothesis.

Measure of both valence and speed

Finally, participants' scores on a measure reflecting simultaneously both valence and speed of responding were analysed. Scores on this measure ranged from -2000 (indicating quick disagreement) to 2000 (indicating quick agreement); less extreme scores reflected less quick responding. With participants' eight scores each on the new measure as dependent variables, I conducted the same mixed-model ANOVA as before. As with the dichotomous judgement data, I will focus exclusively on the

highest-order effect relevant to the repeated assessment of judgements. For a comprehensive discussion of all ANOVA effects, see Appendix C2.

As the highest-order effect of interest, the analysis revealed a three-way interaction of statement valence, onset of first prompt, and ordinal position of response, $F(1,123) = 8.89$, $p < .004$, $MSE = 447434$. See Table 14 for means and standard errors. The interaction was not qualified by higher-order interactions involving cue valence, Fs < 1.4, *n.s.* - To break down the three-way interaction, I tested the simple interactions within each level of statement valence.

A test for the simple interaction within *negative statement valence* revealed main effects of onset of first prompt as well as of ordinal position of response, qualified by an interaction of these factors; $F(1,123) = 6.08$, $p < .02$, $MSE = 333578$ for the interaction. Accordingly, scores for first responses were of a lesser magnitude when the first prompt was presented very early ($M = -849$) than when it was presented not very early ($M = -1341$), $p < .001$ for the pairwise comparison. Scores for second responses showed a similar pattern, $M_{\text{very early}} = -1123$, $M_{\text{not very early}} = -1362$, $p < .02$ for the pairwise comparison. Thus, in line with expectations, very early prompt presentation actually reduced scores on the combined speed-and-valence measure for first responses. Counter to expectation, however, the same effect was observed for second responses (although not as strongly).

A test for the simple interaction within *positive statement valence* showed no main effects of either onset of the first prompt or ordinal position of response, but the factors interacted, $F(1,123) = 4.13$, $p < .05$, $MSE = 472553$. Although scores for first responses were of a somewhat lesser magnitude when the first prompt was presented very early ($M = 673$)

than when it was presented not very early (M = 847), the difference was not significant, $p > .20$ for the pairwise comparison. Scores for second responses also did not differ significantly ($M_{\text{very early}}$ = 865, $M_{\text{not very early}}$ = 791), $p > .60$ for the pairwise comparison. Thus, counter to expectations, very early presentation of the first prompt did not exert an impact on first responses. In line with expectations, an impact on second responses was not found either.

Table 14. Scores on the speed-and-valence measure as a function of statement valence, onset of first prompt, and time of responding (Study 2).

	Statement valence			
	negative		positive	
	First prompt onset		First prompt onset	
	not v.early	v.early	not v.early	v.early
Response signal				
First	−1341 (68)	−849 (67)	847 (98)	673 (97)
Second	−1362 (72)	−1123 (71)	791 (113)	865 (112)

Note. Estimated marginal means on a measure of both speed and valence, in response to first (top row) or second (bottom row) response signals. Standard errors are given in parentheses.
"v.early" = "very early"

To sum up, on the measure that took speed as well as valence of responding simultaneously into account, a significant three-way interaction of statement valence, onset of first prompt, and ordinal position of response was found that had not been observed for dichotomous judgements. The pattern of means underlying it was not always in line with expectations, but may be considered generally supportive (see the Discussion section below). An effect of cue valence on this interaction was not observed.

Discussion

At the core of Study 2 was a persuasion paradigm that manipulated cue valence and statement valence of items orthogonally. Participants responded twice to each item, with the first of these responses being requested either very early or not very early, and the second one being requested late. I expected to find a greater judgemental impact of cue valence on first judgements, and a greater impact of statement valence on second judgements, particularly in conditions where first judgements were requested very early.

Findings regarding the major research questions

The analysis of response latency data showed that first responses occurred later within a window than second responses, but only if first responses were requested very early. Thus, as a means of inducing time pressure, the procedural manipulations (onset of first prompt, and ordinal position of response) were successful. Further, they did not interact with manipulations of item content (cue valence, statement valence, and wording),

suggesting that the technique is fairly insensitive to the particulars of the material processed and may be fruitfully employed with a broad range of contents. It should be noted, however, that the time pressure manipulation did not appear to be overly demanding overall, as indicated by the fact that no-pressure participants had more than 1400 ms left after their decisions on average (which is almost 75 percent of a window's duration), and time-pressure participants had about 1100 ms left (which is more than 50 percent of a window duration).

Participants' dichotomous judgements did, unexpectedly, not reflect the effects of time pressure in the expected way. Instead, judgements were essentially determined by statement valence, and were largely unaffected by the procedural manipulations. In addition, despite a successful manipulation check, no evidence for the expected effects of cue valence was found.

Findings regarding the measure of both speed and valence of responding

Results for the measure of both valence and latency of responding were more promising. Again, no effects of cue valence were found. However, on this measure, statement valence entered into a significant three-way interaction with the procedural manipulations. Specifically, when the first prompt was presented very early (as compared to: not very early), first judgements of negative-valence statements showed a reduced impact of statement valence. An effect into the same direction, although not as strongly, was found for second responses to negative-valence statements. This may indicate a method-induced processing bias. Results for positive-valence statements did not confirm the reduced impact of statement valence

for first responses to early presented first prompts, but also did not confirm the possible method-induced bias.

Despite the unexpected negative-positive asymmetry, and despite the indication of a possible method-induced processing bias I considered findings for the new measure promising. First, different from the dichotomous judgement data, the new measure did show an impact of the procedural manipulations on judgements, suggesting that the effects of time pressure actually had become diluted across response valence and response latency. The measure of both valence and latency appeared less vulnerable to dilution than a measure of judgement valence alone. Second, the judgemental impact of procedural manipulations was into a meaningful direction. Specifically, I had expected to find an increased use of cue valence under conditions of time pressure. This did not occur. Instead, results speak to a reduced use of (negative) statement valence under time pressure, which is exactly what would be expected from a dual-process point of view when no cues are available (or available cues do, for some reason, not work).

The asymmetry of findings between negative and positive statements was not expected. It may tentatively be interpreted as the effect of an initial acquiescence response bias (Knowles & Condon, 1999) that dissipates over time. Accordingly, for statements of a *positive* valence, participants may have tended to agree in their very early first judgements as the result of automatic acceptance, and may have tended to agree in their not very early judgements as the result of more effortful processing. For statements of a *negative* valence, in contrast, participants may have tended to agree automatically in very early first judgements, but may have tended to disagree in not very early first judgements due to effortful reconsideration. Note that for positive

statements, automatic and controlled processes lead to the same judgement, whereas judgements should differ for negative statements. This pattern was actually observed in the present data. However, as the negative-positive asymmetry may well be due to the particular materials used in the study, a conceptual replication using different statements would be desirable before conclusions are drawn.

More of a concern was the indication of a possible method-induced bias found with negative-valence (although not with positive-valence) statements. Again, a replication was called for. Should the bias prove stable across experiments, the usefulness of the within-subjects approach would certainly require reconsideration.

Summary and conclusions

Response latencies were found to be short overall, but to reflect procedural manipulations nevertheless. Dichotomous judgements were found to be basically a function of statement valence, whereas effects of both cue valence and procedural manipulations were largely absent. Effects of procedural manipulations were, however, found on the new measure. The new measure also indicated that a processing bias might be triggered by the instrument.

Taken together, findings called for a modified replication of the study. The absence of cue effects, in the simultaneous presence of successful cue manipulation checks, suggested that a more salient presentation of cues might be useful. Further, response time results suggested that a less lenient timing of response requests was possible. The domination of both first and

second dichotomous judgements by statement valence suggested that a less lenient timing was also desirable. With respect to the negative-positive-asymmetry and to the possible processing bias introduced by the method, no changes were introduced into the follow-up study.

Study 3

Overview

Study 3 was a replication of Study 2, with two modifications. First, I used stricter timing parameters for the onset of prompts, and second, the display of items was changed so as to integrate the two components of an item (cue, and statement) more strongly.

Method

Participants

Twenty-one students of various courses (13 female, 8 male) completed the study in laboratory sessions with one or two participants at a time. They received £2 in return. An additional participant was run, but failed to respond to any out of four second prompts. This participant's data were therefore excluded from analyses.

Materials

The same cues, statements, and item construction procedures as in

Study 2 were used. The only modification concerned the screen display of items. Cues had been displayed in the first line of the display box previously, in order to render them salient. I was made aware that the very same feature (presentation of cues in an extra line) could be used by participants to strategically skip reading of cue information, in order to gain some extra time for evaluating the statements. To prevent such a strategy, cues were followed by statements immediately (that is, in the same line) in Study 3. To further reduce the visual separation of cues and statements, the extra blanks and quotation marks surrounding statements were removed. The colon between cues and statements was, however, retained. In effect, it was now necessary to read the cue information, if only to determine where a statement began.

Procedure and instructions

Procedure and instructions were the same as in Study 2.

Timing parameters

In Study 2, the onset of the very early prompt had been determined by computing the average of predicted reading time and predicted evaluation time, and subtracting one second (50 percent of the window's duration) from this average. Thereby, the *centre* of the window had been made to coincide with the average. By subtracting one additional second in Study 3, the window was shifted such that its *end* (or offset) would coincide with the average of predicted reading and evaluation times. The algorithm was not changed otherwise, which implies that both the not very early and the late

prompt did occur one second sooner, too. Given the previous discussion, there was no reason to correct for this effect.

Design

Like Study 2, Study 3 had a fully factorial, mixed design with three within-subjects factors and two between-subjects factors. The within-subjects factors were ordinal position of response (first vs. second response towards the same item), item cue valence (negative vs. positive), and item statement valence (negative vs. positive). The between-subjects factors were onset of the first out of two response prompts (very early vs. not very early) and item list used (assertions vs. negations).

Due to the small sample size, the assignment of cues to statements could not be fully counterbalanced.

Dependent variables were participants' repeated (first vs. second) agree/disagree-responses towards each of four target items, as well as the associated response latencies; and further, the scores on the new measure.

Hypothesis

I expected to find a threeway-interaction of ordinal position of response, onset of first prompt, and statement valence, such that second responses should reflect statement valence *independent* of the onset of first prompt manipulation, whereas first responses should reflect statement valence *as a function of* onset of first prompt.

More specifically, second responses should always reflect statement valence, whereas first responses should strongly reflect statement valence

only if the first prompt was presented not very early. If it was presented very early, first responses should reflect cue valence. Failing that, a *reduced* impact of statement valence on first responses in very-early-conditions would also be in line with expectations.

Results

Missed responses

Overall, participants missed seven first responses (8.3%), and one second response (1.2%). As before, in terms of absolute frequencies, more misses occurred in response to first (as opposed to second) prompts. I conducted three separate ANOVAs with numbers of missed responses (first, second, and first plus second responses, respectively) as the dependent variables; the between-subjects factors were both item list and onset of the first prompt. These analyses revealed no differences in numbers of missed responses between conditions, Fs < 1.8, *n.s.*

Manipulation checks

Next, I conducted a mixed-model ANOVA with the four manipulation checks as the dependent variables; onset of the first prompt and item list were the between-subjects factors, and cue valence as well as cue gender were the within-subjects factors. As expected, the analysis revealed a main effect of cue valence such that positive cues (M = 4.59) triggered more strongly than negative cues (M = 2.74) the expectation that a not-yet-known statement from a person so described would be agreed with, $F(1,17) = 28.91$,

$p < .001$, $MSE = 2.47$. Thus, like in the previous study, the means of cue manipulation checks were found on expected sides of the response scale. - For higher-order interactions involving cue valence, Fs < 2.6, ps > .13. Different from Study 2, expectations of agreement did not vary as a function of cue gender; for the main effect, $F < 1$; for higher-order interactions, Fs < 2.6, ps > .12.

There was a main effect of onset of first prompt such that, overall, very early prompt onset was associated with less positive cue perception ($M = 3.31$) as compared to not very early prompt onset ($M = 4.03$), $F(1,17) = 7.88$, $p < .02$, $MSE = 1.36$. The effect tended to be more pronounced in conditions where assertions (as opposed to negations) were used, although the interaction was only marginally significant, $F(1,17) = 3.58$, $p < .08$, $MSE = 1.36$. Importantly, however, onset of first prompt did not interact with cue valence; for the two-way interaction, $F < 1$; for higher-order interactions involving onset of first prompt and cue valence, Fs < 2.6, ps > .13. Finally, for the main effect of item list, $F < 1$.

Taken together, results suggested that, by and large, the manipulation of cue valence was successful.

Response times

As in the previous study, I conducted a mixed-model ANOVA with participants' eight response time scores each (representing the time left within a window after a key was pressed) as dependent variables. Onset of first prompt, and item list were the between-subjects factors; cue valence, statement valence, and ordinal position of response were the within-subjects

factors.

The analysis revealed a significant main effect of ordinal position of response, $F(1,17) = 6.84$, $p < .02$, $MSE = 281445$. It was qualified by the expected interaction of this factor with onset of the first prompt, $F(1,17) = 6.71$, $p < .02$, $MSE = 281445$. Accordingly, for *first responses*, less time was left after a decision when the first prompt appeared very early ($M = 1063$ ms) than when it appeared not very early ($M = 1288$ ms). The difference was marginally significant, $p < .10$ for the pairwise comparison. For *second responses*, the opposite pattern was observed, with more time remaining when the first prompt had been shown very early ($M = 1491$ ms) than when it had been shown not very early ($M = 1290$ ms). The difference was significant, $p < .05$ for the pairwise comparison. Thus, whereas the expected two-way interaction did actually occur, the pattern of means underlying it was not in line with expectations. Specifically, first responses were affected by onset of the first prompt in the expected direction, but only marginally so. Second responses should not have been affected by onset of the first prompt, but actually were affected, and even significantly so.

The interaction just described tended to be qualified by a marginally significant five-way interaction of all factors, $F(1,17) = 3.31$, $p < .09$, $MSE = 162523$. Item list further entered into two marginally significant interactions, firstly with onset of the first prompt, $F(1,17) = 4.08$, $p < .06$, $MSE = 236542$, and secondly, with both statement valence and onset of the first prompt, $F(1,17) = 3.59$, $p < .08$, $MSE = 135353$. No other effects occurred, Fs < 2.2, ps > .16. Given the converging evidence of three marginally significant interactions pointing to an effect of item list, it appeared appropriate to look at response times separately for assertions versus

negations.

Response times for assertions. I repeated the mixed-model ANOVA described above, using only data from participants who had been exposed to assertions, thus omitting the factor item list. The interaction of ordinal position of response with onset of the first prompt was found to be marginally significant, $F(1,8) = 4.96$, $p < .06$, $MSE = 444043$. No other effects were found, Fs < 2.3, ps > .16. - For *first responses*, less time was remaining after a decision when the first prompt had been shown very early ($M = 927$ ms) than when it had been shown not very early ($M = 1423$ ms). The difference was significant, $p < .04$ for the pairwise comparison. After *second responses*, in contrast, about the same time was left within a window, independent of onset of the first prompt ($M_{very\ early} = 1427$ ms, $M_{not\ very\ early} = 1259$ ms), $p > .30$ for the pairwise comparison. Thus, for assertions, onset of the first prompt selectively affected first (but not second) responses, and did so in the expected direction.

Response times for negations. The same mixed-model ANOVA revealed different results when using the negation data. Here, response times were determined by ordinal position of response such that less time was left after first responses ($M = 1177$ ms) than after second responses ($M = 1438$ ms), $F(1,9) = 10.92$, $p < .01$, $MSE = 136913$. Importantly, onset of first prompt did not exert a main effect, $F < 1.8$, *n.s.*, and did not interact with ordinal position of response, $F < 1.4$, *n.s.* For other effects, Fs < 3.2, ps > .10. Thus, for negations, onset of the first prompt left both first and second responses unaffected.

To sum up, for *assertions*, I found the expected pattern whereby very early (as opposed to: not very early) onset of the first prompt should

decrease speed of responding for first responses, but not for second responses. *Negations*, in contrast, were not affected by manipulations of the onset of the first prompt. Disregarding the difference between assertions and negations, response time scores reflected manipulations of process variables only (i.e., manipulations of onset of first prompt, and of number of response), but not of statement content (i.e., statement valence, and cue valence).

Dichotomous judgements of agreement

Response time data had suggested that participants processed assertions somewhat differently than negations. Therefore, it appeared appropriate to look at the judgement data separately for assertions and negations.

Agreement judgements for assertions. The mixed-model ANOVA of participants' eight judgements each was repeated, using only the data from participants who had been exposed to assertions, thus omitting the item list factor. The analysis showed a main effect of statement valence, $F(1,8) = 160.17$, $p < .001$, $MSE = .30$. This was qualified by an interaction with onset of first prompt, $F(1,8) = 6.00$, $p < .05$, $MSE = .30$. The two-way interaction, in turn, was qualified by a marginally significant three-way interaction of these factors with ordinal position of response, $F(1,8) = 4.50$, $p < .07$, $MSE = .10$. Other effects did not occur, Fs < 3.3, ps > .11. - I expected to find a three-way interaction of onset of first prompt and ordinal position of response with statement valence such that, within each level of statement valence, first responses should differ as a function of prompt onset, whereas second responses should not differ. As the three-way

interaction that actually occurred showed (see Table 15 for means), the expected pattern was found for statements of positive valence (for the pairwise comparison of first responses, $p < .05$; for the pairwise comparison of second responses, $p > .20$). For statements of negative valence, the pattern of means was consistent with the expectation, but the difference between first responses failed to reach significance (for the pairwise comparison of first responses, $p < .18$; for the pairwise comparison of second responses, $p > .99$). This pattern suggested that the three-way interaction may have been due essentially to judgements of positive-valence (but not negative-valence) statements. In order to follow up this possibility, I conducted tests for the simple interactions within levels of statement valence. These analyses revealed no significant interactions involving prompt onset and ordinal position of response for either positive-valence or negative-valence statements, Fs < 1.9, ps > .20. Hence, responses to both kinds of statements contributed to the marginally significant three-way interaction just described. Thus, by and large, the pattern of means for judgements of assertions was in line with expectations.

Table 15. Dichotomous judgements of assertions as a function of statement valence, onset of first prompt, and time of responding (Study 3).

	Statement valence			
	negative		positive	
	First prompt onset		First prompt onset	
	not v.early	v.early	not v.early	v.early
Response signal				
First	−1.00 (.14)	−.70 (.14)	1.00 (.17)	.40 (.17)
Second	−.80 (.20)	−.80 (.20)	.90 (.19)	.60 (.19)

Note. Estimated marginal means for dichotomous judgements of disagreement or agreement in response to first (top row) or second (bottom row) response signals. Standard errors are given in parentheses.
"v.early" = "very early"

Agreement judgements for negations. Results were different when analysing the negation data only. A main effect of statement valence tended to be qualified by an interaction with ordinal position of response, such that second responses (as compared to first responses) reflected statement valence somewhat more strongly (for positive-valence statements: $M_{first\ response}$ = .11, $M_{second\ response}$ = .20, p > .50 for the pairwise comparison; for negative-valence statements: $M_{first\ response}$ = -.54, $M_{second\ response}$ = -.90, p < .12

for the pairwise comparison). For the statement valence main effect, $F(1,9) = 28.64$, $p < .001$, $MSE = .58$; for the interaction, $F(1,9) = 3.49$, $p < .10$, $MSE = .32$. Importantly, the interaction of statement valence and number of response was not qualified by higher-order interactions with onset of first prompt, Fs < 1. - Finally, a three-way interaction of onset of first prompt, ordinal position of response, and cue valence was observed, $F(1,9) = 7.10$, $p < .03$, $MSE = .11$. No other effects occurred, Fs < 2.3, ps > .16. When following up the three-way interaction by running separate analyses within levels of prompt onset, I found an interaction of cue valence and ordinal position of response when the first prompt was presented not very early, but no effects of these factors when the first prompt was presented very early. Specifically, for *not very early first prompts*, participants' first responses were in line with cue valence ($M_{positive\ cue} = 0$, $M_{negative\ cue} = -.200$, $p > .40$ for the pairwise comparison), whereas their second responses were opposite to cue valence ($M_{positive\ cue} = -.400$, $M_{negative\ cue} = 0$, $p < .18$ for the pairwise comparison), $F(1,4) = 10.29$, $p < .04$, $MSE = .09$. For *very early first prompts*, in contrast, Fs < 1 for both main effects and the interaction of cue valence and time of response. - Thus, the expected three-way interaction was not found for negations; specifically, interactive effects of statement valence and number of response were not affected by the onset of first prompt manipulation. If the onset manipulation affected anything, then late onset of the first prompt seemed to trigger an unexpected contrast effect away from cue valence in participants' second responses.

To sum up, similar to the latency data, participants' dichotomous judgements were in line with expectations for assertions (although the

predicted interaction became only marginally significant), but not for negations.

Combined measure of both speed and valence

The analyses of both response time data and dichotomous judgement data had shown different results when looking at assertions versus negations separately. Therefore, it seemed appropriate to separately look at these data for the speed-and-accuracy measure, too.

Speed-and-valence scores for assertions. First, I conducted a mixed-model ANOVA with the eight scores on the new measure as dependent variables, using all experimental factors except for item list, and using only data from participants who had been exposed to assertions. The analysis revealed a main effect of statement valence, $F(1,8) = 166.58$, $p < .001$, $MSE = 560349$. The main effect was qualified by a two-way interaction with onset of first prompt, $F(1,8) = 5.43$, $p < .05$, $MSE = 560349$. The two-way interaction, in turn, was qualified by a three-way interaction of statement valence, onset of first prompt, and ordinal position of response, $F(1,8) = 6.12$, $p < .04$, $MSE = 431462$. No other effects were observed, Fs < 2.1, ps > .19. - As outlined in the judgement data section for assertions, I expected a specific three-way interaction for these data. Compared to the dichotomous judgements, scores on the measure that simultaneously took speed and accuracy of responding into account corresponded more closely to the expected pattern of means (see Table 16 for means). Specifically, in the judgement data, *first responses* differed as a function of onset of first prompt for positive-valence statements, but not for negative valence statements.

With the new measure, the difference for positive statements was confirmed (for the pairwise comparison, $p < .002$). For negative statements, the previously non-significant difference was now found to be marginally significant (for the pairwise comparison, $p < .09$). At the same time, *second responses* still did not differ significantly (ps > .60 for the pairwise comparisons). - Finally, tests for the simple interactions within levels of statement valence revealed a significant interaction of onset of first prompt and ordinal position of response for positive-valence statements, $F(1,8) = 6.22$, $p < .04$, $MSE = 207144$. For negative-valence statements, the interaction tended to approach marginal significance, $F(1,8) = 3.30$, $p < .11$, $MSE = 409158$. Taken together, for assertions, scores on the new measure were even more strongly in line with the hypothesis than judgement data had been.

Table 16. Scores for assertions on the speed-and-valence measure as a function of statement valence, onset of first prompt, and time of responding (Study 3).

	Statement valence			
	negative		positive	
	First prompt onset		First prompt onset	
	not v.early	v.early	not v.early	v.early
Response signal				
First	−1458 (221)	−850 (221)	1387 (123)	489 (123)
Second	−1097 (281)	−1223 (281)	1159 (297)	979 (297)

Note. Estimated marginal means on a measure of both speed and valence, in response to first (top row) or second (bottom row) response signals. Standard errors are given in parentheses.
"v.early" = "very early"

Speed-and-valence scores for negations. The same mixed-model ANOVA as before was conducted, this time using only the data from participants who had been exposed to negations. The analysis revealed a main effect of statement valence that tended to be qualified by an interaction with ordinal position of response. Similar to the judgement data, second responses (as compared to first responses) reflected statement valence

somewhat more strongly (for positive statements: $M_{first\ response}$ = 122, $M_{second\ response}$ = 280, *n.s.*; for negative statements: $M_{first\ response}$ = -862, $M_{second\ response}$ = -1422, for the pairwise comparison, $p < .09$). For the statement valence main effect, $F(1,9) = 32.18$, $p < .001$, $MSE = 1222875$; for the interaction, $F(1,9) = 3.54$, $p < .10$, $MSE = 793605$. Finally, a three-way interaction of onset of first prompt, ordinal position of response, and cue valence occurred, $F(1,9) = 7.24$, $p < .03$, $MSE = 161832$. No other effects were observed, Fs < 3.1, ps > .11. - When following up the three-way interaction by running separate analyses within levels of prompt onset, I found an interaction of cue valence and ordinal position of response when the first prompt was presented not very early, but no effects of these factors when the first prompt was presented very early. Specifically, for *not very early first prompts*, participants' first responses were in line with cue valence ($M_{positive\ cue}$ = -56, $M_{negative\ cue}$ = -295, $p > .40$ for the pairwise comparison), whereas their second responses were opposite to cue valence ($M_{positive\ cue}$ = -629, $M_{negative\ cue}$ = -103, $p < .12$ for the pairwise comparison), $F(1,4) = 28.83$, $p < .01$, $MSE = 50884$. For *very early first prompts*, in contrast, Fs < 1 for both main effects and the interaction of cue valence and time of response. - Thus, like the dichotomous judgement data, scores on the new measure were not in line with expectations.

To sum up, the predicted three-way interaction was found for assertions, but not for negations. Different from the dichotomous judgement data, the interaction was significant when using scores on the measure simultaneously reflecting both speed and valence of responses as the dependent variable.

Discussion

Compared to Study 2, the present study adopted a stricter timing of response windows, as well as a different display of cues and statements. Again, the expected effects of cue valence were not observed. Nevertheless, results supported predictions, at least for assertively phrased items.

Assertions. Response latencies and dichotomous judgement data showed the expected interactions. The interactions were only marginally significant, but means differed into the predicted directions. With the measure of both speed and valence, the hypothesised three-way interaction reached significance. Scores on the measure generally reflected statement valence. Within levels of statement valence, second responses did not differ as a function of onset of first prompt. First responses, in contrast, did differ depending on prompt onset. Specifically, very early (as compared to: not very early) prompt onset reduced the extremity of scores. Although the reduction was only marginally significant for negative-valence statements, the overall significant three-way interaction supported the experimental hypothesis. Thus, the instrument did not seem to introduce a method-induced processing bias.

Negations. Both response latencies and judgements suggest that participants processed negations in a different way than assertions. Specifically, participants required more time for first than for second responses, independent of the onset of the first prompt. Note that the systematic difference in word numbers between assertions and negations (i.e., the additional word "not") had been accounted for by the online algorithm. The insensitivity of latencies for first responses to the onset of prompt manipulation therefore suggests that the comprehension of negations

requires a cognitive operation that is not involved in the comprehension of assertions - over and above the time needed to read one word more. With respect to judgments, participants seemed to consider cue valence in their *second* responses, an effect that was not expected, and, more importantly, that is unlikely to be the result of automatic processing. Together, these findings suggest that timing parameters were appropriate for assertions, but were too strict to allow for a comprehensive processing of negations.

Summary and Discussion

The present chapter introduced an instrument for observing the construction of a social judgement within-subjects. The instrument was explored in two studies. Adopting a persuasion paradigm, Study 2 aimed to demonstrate a) that participants switch between agreement and disagreement in a predictable order, and b) that the instrument does not introduce a processing bias. Initial evidence was provided, but the specific hypothesis tested was not fully supported. Study 3 used stricter timing parameters than Study 2. Results confirmed the hypothesis for assertions, but not for negations. Specifically, for assertions, requesting a response very early (as compared to late) reduced the judgemental impact of statement valence. At the same time, late responses were shown not to be impacted by the overt reporting of a preliminary judgement, excluding the possibility of a method-induced processing bias.

Two findings of these studies deserve closer attention. First, why did findings differ between assertions and negations? And second, why were no cue effects observed?

With respect to the assertion-negation asymmetry, it is useful to recall the response latency results from Studies 2 and 3. In Study 2, the procedural manipulations (onset of first prompt, and ordinal position of response) interactively affected assertions in the same way as negations. In Study 3, latencies for assertions were affected by the same interaction, whereas latencies for negations were affected only by ordinal position of response. A straightforward explanation of this pattern holds that the timing parameters in Study 2 were sufficiently lenient to allow for the high-effort processing of both assertions and negations, whereas the parameters of Study 3 were too strict to allow for a sufficient amount of high-effort processing to occur with the more difficult statement category, negations. If so, then future studies should be able to extend the assertion findings to negations by providing more time, for instance by assessing reading and evaluation times separately for the two statement types, and using separate speed indices in subsequent steps.

The question why there were no cue effects in Studies 2 and 3 is not easy to answer. The cues had been pretested successfully under conditions of time pressure. Further, cue manipulation checks at the end of each study were in line with expectations. Nevertheless, cues did not exert the predicted impact on judgements. Note, however, that in both pretest and manipulation check participants only *confirmed* that the cues would trigger a given experimenter-provided expectation - they did not have to generate the expectation themselves. Gilbert and Hixon (1991) found that persons under cognitive load did use an available stereotype if it was already activated, but did not spontaneously activate a stereotype that was available. Possibly, beyond being made salient, cues would have required to be activated in the present studies in order to exert an impact. This should be investigated in

future research using the instrument by asking participants to complete a pre-experimental questionnaire (similar to the pretest study) that does versus does not comprise the to-be-used target cues.

Empirical clarification of these questions appears highly desirable as, obviously, an instrument that works only with assertions is less useful than one that may be applied to both assertions and negations. Similarly, the present research has not shown that cues determine early stages of the judgement construction process, whereas statements determine late stages. Instead, it has shown that the impact of statements is smaller at early than at late stages. Because this finding may be explained by "regression towards the mean", a demonstration of cue effects at an early, but not at a late stage would have been preferable. It has however been shown how to assess judgements repeatedly, without introducing a method-induced bias, both at and before the point in time where high-effort processing exerts its greatest judgemental impact.

Summary and conclusions

The present work compared classic and connectionist approaches to social judgement. Chapter 1 dealt with the domain of person perception. Two classic approaches and one connectionist account of attribute emergence were pinpointed against each other. Each of these models can explain the emergence of novel attributes. For the comparison, I derived the models' predictions for a different variable, namely response times. Results for this variable seemed to suggest that there may actually be two distinct mechanisms of attribute emergence, one of them being more automatic, and the other being more consciously controlled. Such a distinction had not been visible in the proportions of emergent attributes. Pending replication of the result pattern, we may tentatively conclude that the connectionist model contributed to our understanding of attribute emergence by providing a mechanism whereby emergent attributes are inferred quickly. More importantly, however, the connectionist model explained attribute emergence by preconscious conceptual combination. That mechanism may provide an explanation of emergent attributes that occur if category combinations are congruent and unsurprising - from the classic approaches' point of view, emergent attributes should not be observed to begin with.

Chapter 2 dealt with persuasive communication. Based on assumptions common to three classic models of persuasion, I proposed a spreading activation model of persuasion. Essentially, the connectionist model embodied the assumption that persuasion by arguments, but not persuasion by cues, requires the mediation by cognitive responses.

Simulations of actual persuasion studies from the literature confirmed that the connectionist account can accomodate the patterns of data from human participants collected in these studies equally well as the classic accounts. However, the classic accounts featured additional process assumptions that were not required to reproduce the data patterns. The proposed network model's cognitive mechanism relied strongly on preliminary, cue-based attitude judgements as mediators of final attitude judgements. If so, then attitudes that reflect cue valence should be observable, within-subjects, even if persons process a persuasive communication under conditions of high effort and motivation, and will ultimately arrive at an attitude judgement of the opposite valence. To demonstrate the effect empirically, it would be necessary to assess both preliminary and final attitude judgements within-subjects.

First steps toward an instrument that allows for the repeated assessment of judgements within-subjects were presented in Chapter 3. Specifically, the instrument predicts the point in time where people would spontaneously provide a judgement of a stimulus. Using response prompts, researchers gain the opportunity of requesting a judgements before that point in time, afterwards, or repeatedly. Results of two studies using the instrument suggested that the timing of response prompts, as determined by an online algorithm, actually corresponds to points in time before versus after the time where a participant would spontaneously choose to respond. The studies further investigated the question if the provision of the first judgement exerts an undue influence on the second judgement. Overall, results were in line with the view that this is not the case. The instrument was found to reflect particulars of stimulus statements' wording (assertion vs. negation), a result

that should be taken into account when selecting stimulus materials. Further, responses requested before a participant would respond spontaneously were found not to be affected by easy-to-process cue information, a result counter to expectations. I proposed a follow-up study that may clarify if the application of cues in the paradigm possibly depends on their prior activation.

Taken together, the present chapters suggest that a connectionist approach to social judgement may complement classic approaches (Chapter 1), may provide a viable alternative explanation of a domain's basic findings (Chapter 2), and may stimulate the development of new research techniques (Chapter 3).

References

Asch, S. E., & Zukier, H. (1984). Thinking about persons. *Journal of Personality and Social Psychology, 46,* 1230-1240.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173-1182.

Bohner, G., Crow, K., Erb, H.-P., & Schwarz, N. (1992). Affect and persuasion: Mood effects on the processing of message content and context cues and on subsequent bevaviour. *European Journal of Social Psychology, 22,* 511-530.

Bohner, G., & Siebler, F. (1999). Paradigms, processes, parsimony, and predictive power: Arguments for a generic dual-process model. *Psychological Inquiry, 10,* 113-118.

Chaiken, S, Duckworth, K. L.,. & Darke, P. (1999). When parsimony fails. *Psychological Inquiry, 10,* 118-123.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 212-252). New York: Guilford Press.

Chaiken, S., & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: Effects of source credibility, argument ambiguity, and task importance on attitude judgment. *Journal of Personality and Social Psychology, 66,* 460-473.

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

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

Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247-282). Hillsdale, NJ: Lawrence Erlbaum Associates.

Fiedler, K. (1996). Explaining and simulating judgment biases as an aggregation phenomenon in probabilistic, multiple-cue environments. *Psychological Review, 103,* 193-214.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46,* 107-119.

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60,* 509-517.

Giner-Sorolla, R., & Chaiken, S. (1997). Selective use of heuristic and systematic processing under defense motivation. *Personality and Social Psychology Bulletin, 23,* 84-97.

Greenwald, A. G. (1968). Cognitive learning, cognitive response to persuasion, and attitude change. In A. G. Greenwald, T. C. Brock, & T. M. Ostrom (Eds.), *Psychological foundations of attitudes* (pp. 147-170). New York: Academic Press.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test.

Journal of Personality and Social Psychology, 74, 1464-1480.

Hastie, R., Schroeder, C., & Weber, R. (1990). Creating complex social conjunction categories from simple categories. *Bulletin of the Psychonomic Society, 28,* 242-247.

Howard, D. J. (1997). Familiar phrases as peripheral persuasion cues. *Journal of Experimental Social Psychology, 33,* 231-243.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513-541.

Judd, C. M., & Kulik, J. A. (1980). Schematic effects of social attitudes on information processing and recall. *Journal of Personality and Social Psychology, 38,* 569-578.

Knowles, E. S., & Condon, C. A. (1999). Why people say "yes": A dual-process theory of acquiescence. *Journal of Personality and Social Psychology, 77,* 379-386.

Kruglanski, A. W., & Thompson, E. P. (1999a). Persuasion by a single route: A view from the unimodel. *Psychological Inquiry, 10,* 83-109.

Kruglanski, A. W., & Thompson, E. P. (1999b). The illusory second mode or, the cue is the message. *Psychological Inquiry, 10,* 182-193.

Kunda, Z., Miller, D. T., & Claire, T. (1990). Combining social concepts: The role of causal reasoning. *Cognitive Science, 14,* 551-577.

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review, 103,* 284-308.

Maheswaran, D., & Chaiken, S. (1991). Promoting systematic processing in low-motivation settings: Effect of incongruent information on

processing and judgment. *Journal of Personality and Social Psychology, 61,* 13-25.

McClelland, J. L., & Rumelhart, D. E. (1986). A distributed model of human learning and memory. In J. L. McClelland & D. E. Rumelhart (Eds.): *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2, pp. 170-215). Cambridge, MA: MIT Press.

Mosler, H.-J., Schwarz, K., Ammann, F., & Gutscher, H. (2001). Computer simulation as a method of further developing a theory: Simulating the Elaboration Likelihood Model. *Personality and Social Psychology Review, 5,* 201-215.

Petty, R. E., & Cacioppo, J. T. (1984). The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology, 46,* 69-81.

Petty, R. E., & Caccioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change.* New York: Springer-Verlag.

Petty, R. E., Cacioppo, J. T., Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology, 41,* 847-855.

Petty, R. E., & Wegener, D. T. (1999). The elaboration likelihood model: Current status and controversies. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 41-72). New York: The Guilford Press.

Petty, R. E., Wheeler, S. C., & Bizer, G. Y. (1999). Is there one persuasion process or more? Lumping versus splitting in attitude change theories.

Psychological Inquiry, 10, 83-109.

Read, S. J., & Marcus-Newhall, A. (1993). Explanatory coherence in social explanations: A parallel distributed processing account. *Journal of Personality and Social Psychology, 65,* 429-447.

Read, S. J., & Miller, L. C. (1993). Rapist or "regular guy": Explanatory coherence in the construction of mental models of others. *Personality and Social Psychology Bulletin, 19,* 526-541.

Read, S. J., Vanman, E. J., & Miller, L. C. (1997). Connectionism, parallel constraint satisfaction processes, and gestalt principles: (Re)Introducing cognitive dynamics to social psychology. *Personality and Social Psychology Review, 1,* 26-53.

Romero, A. A., Agnew, C. R., & Insko, C. A. (1996). The cognitive mediation hypothesis revisited: An empirical response to methodological and theoretical criticism. *Personality and Social Psychology Bulletin, 22,* 651-665.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119,* 3-22.

Smith, E. R. (1994). Procedural knowledge and processing strategies in social cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 99-151). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Smith, E. R. (1996). What do connectionism and social psychology offer each other? *Journal of Personality and Social Psychology, 70,* 893-912.

Smith, E. R., & DeCoster, J. (1998). Knowledge acquisition, accessibility, and use in person perception and stereotyping: Simulation with a recurrent connectionist network. *Journal of Personality and Social Psychology,*

74, 21-35.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review, 4,* 108-131.

Smith, E. R., & Henry, S. (1996). An in-group becomes part of the Self: Response time evidence. *Personality and Social Psychology Bulletin, 22,* 635-642.

Smith, S. M., & Petty, R. E. (1996). Message framing and persuasion: A message processing analysis. *Personality and Social Psychology Bulletin, 22,* 257-268.

Smolensky, P. (1988). On the proper treatment of connectionism. *Behavioral and Brain Sciences, 11,* 1-74.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences, 12,* 435-502.

Van Overwalle, F. (1998). Causal explanation as constraint satisfaction: A critique and a feedforward connectionist alternative. *Journal of Personality and Social Psychology, 74,* 312-328.

Van Overwalle, F., Labiouse, C., & French, R. (2001). *Connectionist exploration in social cognition.* Unpublished manuscript.

Wänke, M., Bless, H., & Biller, B. (1996). Subjective experience versus content of information in the construction of attitude judgments. *Personality and Social Psychology Bulletin, 22,* 1105-1113.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107,* 101-126.

Appendix

Appendix A

This Appendix comprises a discussion of why the results of Study 1 suggest that two distinct mechanisms of attribute emergence may be at work. Additional data analyses are provided in A1; possible theoretical mechanisms are discussed in A2.

A1. Mediation analyses (Study 1)

According to Kunda, Miller, & Claire's (1990) explanation of attribute emergence, persons experience surprise when encountering a member of a puzzling category combination. Novel attributes emerge as the result of an attributional activity, namely the generation of a causal account. Thus, in Kunda et al.'s model, the construction of a causal account takes the role of a variable intervening between surprise and attribute emergence, i.e., of a mediator. The ANOVA results reported in Chapter 1 already suggest that time-consuming cognitive activity (like the construction of a causal account) may be a sufficient, but not a necessary factor contributing to the emergence of novel attributes. In addition, I conducted formal mediation analyses. It should be noted that these are post-hoc analyses as they rely on a difference between target categories that had been found empirically, but had not been predicted.

Separate mediation analyses were run for the two target categories ("nurse", and "mechanic"). In the analyses, I used proportions of emergent attributes as the dependent variables. Self-reports of experienced surprise

(or: manipulation check data) served as the independent variables, because they were more fine-grained than the dichotomous variable "experimental condition". Overall response times were assumed to reflect the amount of cognitive activity triggered by experienced surprise; they served as the mediator variables. Following the procedure described in Baron and Kenny (1986), I estimated three regression equations by "first, regressing the mediator on the independent variable; second, regressing the dependent variable on the independent variable; and third, regressing the dependent variable on both the independent variable and on the mediator" (p. 1177). For informative purposes, a fourth equation was estimated where I regressed the dependent variable on the mediator alone. Mediation may be assumed under the following conditions: "First, the independent variable must affect the mediator in the first equation; second, the independent variable must be shown to affect the dependent variable in the second equation; and third, the mediator must affect the dependent variable in the third equation. If these conditions all hold in the predicted direction, then the effect of the independent variable on the dependent variable must be less in the third equation than in the second" (p. 1177)

Results for the target category "Nurse". As can be seen from Figure 6, greater experienced surprise resulted in a greater proportion of emergent attributes, $\beta = .41$, $p < .01$ (second equation). In addition to this direct effect, surprise tended to exert an indirect effect on the proportion of emergent attributes. Specifically, greater surprise triggered more cognitive activity (as assessed by the overall response time), $\beta = .33$, $p < .03$ (first equation); more cognitive activity, in turn, tended to increase the proportion of emergent attributes, $\beta = .25$, $p < .10$ (fourth equation). When regressing the proportion

of emergent attributes on both predictors simultaneously (third equation), however, the marginal effect of response time became insignificant, *ß* = .13, *n.s.* Experienced surprise, in contrast, remained an excellent predictor of attribute emergence, *ß* = .36, *p* < .02. Thus, for the target category "nurse", there was no indication of cognitive activity (as assessed by response time) mediating the effect of experienced surprise on the proportion of emergent attributes.

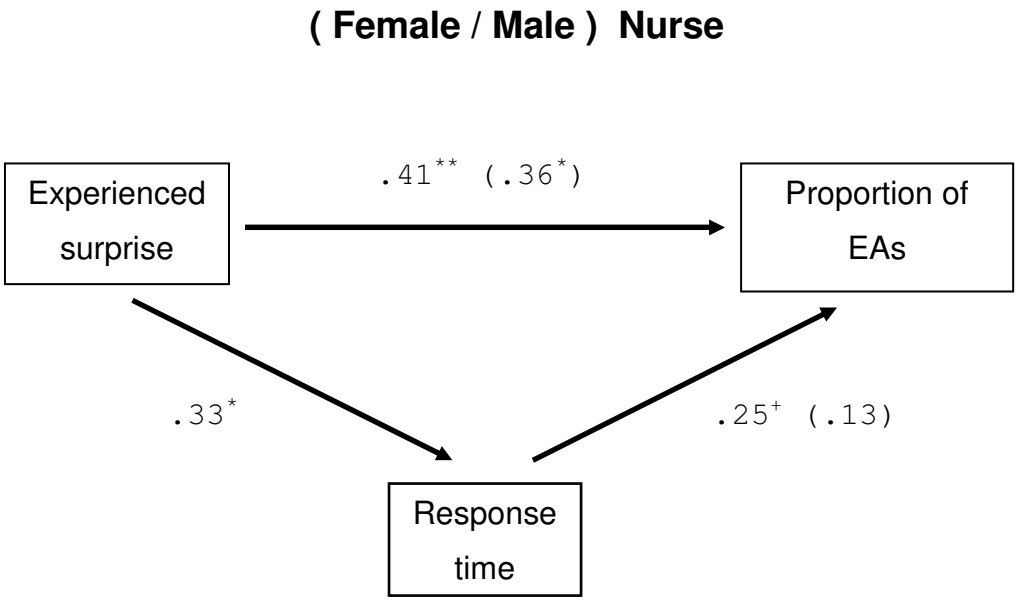

Figure 6. Target person "Nurse": The direct effect of experienced surprise on the proportion of emergent attributes is not mediated by overall response time (Study 1).

Note. Figures are regression beta weights. Weights inside (outside) parentheses are (are not) corrected for the effect of the respective other predictor. EAs: Emergent attributes.

*** p < .01 * p < .05 + p < .10*

Results for the target category "Mechanic". As Figure 7 shows, greater

experienced surprise resulted in a greater proportion of emergent attributes, $\beta = .30$, $p < .05$ (second equation). In addition to this direct effect, surprise exerted an indirect effect on the proportion of emergent attributes. Specifically, greater surprise triggered more cognitive activity, $\beta = .50$, $p < .001$ (first equation); more cognitive activity, in turn, increased the proportion of emergent attributes, $\beta = .36$, $p < .02$ (fourth equation). When regressing the proportion of emergent attributes on both predictors simultaneously (third equation), the effect of response time was reduced, but remained marginally significant, $\beta = .28$, $p < .10$. The effect of experienced surprise on the dependent variable, in contrast, was greatly reduced, $\beta = .16$, *n.s.* Thus, for the target category "mechanic", cognitive activity (as assessed by response time) mediated the effect of experienced surprise on the proportion of emergent attributes.

(Male / Female) Mechanic

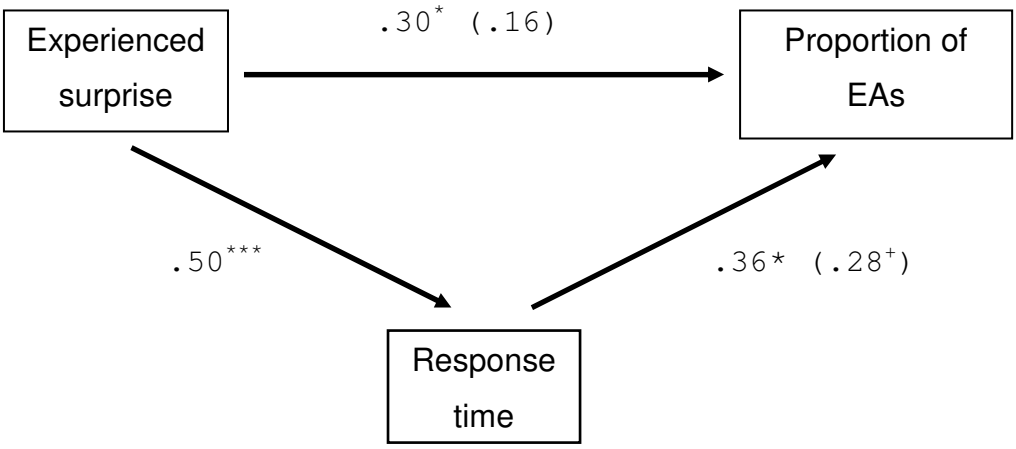

Figure 7. Target person "Mechanic": The direct effect of experienced surprise on the proportion of emergent attributes is mediated by overall response time (Study 1).

Note. Figures are regression beta weights. Weights inside (outside) parentheses are (are not) corrected for the effect of the respective other predictor. EAs: Emergent attributes.

**** p < .001 * p < .05 + p < .10*

To sum up, cognitive activity (as assessed by overall response time) mediated the effect of experienced surprise on the proportion of novel attributes for a target person described as a (male or female) "mechanic", but not for a target person described as a (female or male) "nurse". Like the ANOVA results reported in Chapter 1, mediation analyses were in line with the assumption that novel attributes may be inferred by each of two distinct mechanisms, one of them involving time-consuming amounts of thought, and

the other one being more automatic.

A2. Theoretical mechanisms of attribute emergence (Study 1)

The data from Study 1 support the notion that novel attributes may stem from two distinct processing mechanisms. How does the distinction of two mechanisms map onto the distinction of classical explanations *versus* a connectionist account of attribute emergence? In Chapter 1, I emphasised the classic approaches' focus on cognitively demanding processes, as well as the connectionist model's ability to explain the same empirical phenomenon without assuming the expenditure of cognitive effort. Actually, the division is not quite as clear-cut. For instance, Hastie, Schroeder, and Weber (1990) reported various strategies used by their participants to resolve conflicting implications of dual category membership, among them the identification of a relevant exemplar in long-term memory, or the application of general rules that come to mind (see Chapter 1). In a related vein, Kunda and Thagard (1996) proposed that simultaneous membership in apparently incongruent categories may lead to the activation of a relatively atypical subtype of a stereotype – for instance, a Harvard-educated carpenter might be perceived as a "Harvard-educated hippie, who is believed to be non-materialistic" (p. 299). Because exemplars or subtypes may come to mind without much effort expenditure, classic approaches may well account for emergent attributes that take little processing time to infer. Conversely, the connectionist account may be extended to explain why the inference of a novel attribute sometimes takes an additional amount of processing time. Recall that Smith and DeCoster's (1998) model of attribute emergence relied

on immediately overlapping representations of stereotypes, such that only few steps of pattern completion were required to infer the emergent attribute. However, more complex scenarios are conceivable that require a considerably greater number of intermediate knowledge structures to be activated before an emergent attribute is inferred. Whereas a single step of pattern completion may take place almost instantly, the aggregated time demands of many steps may well add up to a measurable delay. Pending a demonstration of that effect, the connectionist model provides an equally comprehensive framework as classical approaches, with the added advantages outlined in the Introduction section.

Appendix B

This Appendix provides a table showing those pairs of opinion statements (out of 24) that comprised members of opposite valence, as found in the pretest for Study 2.

Table 17. Statement pairs (Pretest for Study 2)

Pair	Statement	t(12)
1(−)	It is better to give directly to beggars rather than to charities.	−3.09
1(+)	It is better to give to charities rather than directly to beggars.	3.43
2(−)	If a friend came into financial difficulty you should refuse to lend them any money.	−6.25
2(+)	If a friend came into financial difficulty you should offer to lend them some money.	7.67
3(−)	Children before age 16 should not be allowed to take first-aid classes.	−6.40
3(+)	Children between 12 and 16 should have to take first-aid classes.	6.74
4(−)	Expensive jewellery is essential to improve one's appearance.	−4.67
4(+)	Expensive jewellery is not essential to improve one's appearance.	4.90

(Table continued on next page)

(Table 17 continues)

5(−)	If people wish to have children, they should go ahead regardless of whether they can financially provide for them.	−6.88
5(+)	If people wish to have children, they should first ensure that they can financially provide for them.	6.44
6(−)	It's not wrong to write and spread computer viruses, as long as no harm is intended.	−11.08
6(+)	It is wrong to write and spread computer viruses, even if no harm is intended.	5.67
7(−)	People who live in areas at risk for natural disasters should not be helped by the government when a disaster strikes.	−18.62
7(+)	People who live in areas at risk for natural disasters should be helped by the government when a disaster strikes.	9.73

Note. Plus- and minus-signs in parentheses depict statement valence (plus: agreeable, minus: not agreeable). T-values stem from one-sample t-tests for difference from the midpoint (4) of the agreement scale ranging from 1 ("not at all") to 7 ("very much").

Appendix C

This Appendix provides a comprehensive discussion of mixed-model ANOVA results for dichotomous judgements of agreement (C1) as well as the measure of both valence and speed (C2) for Study 2.

C1. Dichotomous judgements of agreement (Study 2)

Participants' eight dichotomous judgement scores each were submitted to a mixed-model ANOVA. The between-subjects factors were onset of first prompt (not very early vs. very early), and item list (negations vs. assertions). Cue valence (negative vs. positive), statement valence (negative vs. positive), and ordinal position of response (first vs. second) were the within-subjects factors.

Three main effects were found in the data. The most prominent effect was a main effect of statement valence whereby statements of a positive valence led to greater agreement (M = .56) than statements of a negative valence (M = -.81), $F(1,123) = 567.71$, $p < .001$, MSE = .84. This confirms that participants perceived statement valence in line with pretest results. Further, list of items used (negations vs. assertions) exerted a main effect such that negations led to greater disagreement overall (M = -.25) than assertions (M = -.004), $F(1,123) = 22.36$, $p < .001$, MSE = .68. This reflects the assertion-negation bias described by Knowles and Condon (1999). Finally, when the first prompt was presented very early, less disagreement was observed overall (M = -.07) than when it was presented only later (M = -.18), $F(1,123) = 4.82$, $p < .04$, MSE = .68 (see below for a qualifying

interaction and its discussion).

Interestingly, item list interacted with statement valence. Specifically, positive-valence negations led to average judgements of a lesser magnitude (M = .31) than all other items (Ms for positive-valence assertions, negative-valence assertions, and negative-valence negations were: .81, -.82, and -.81, respectively). For the interaction, $F(1,123) = 18.96$, $p < .001$, MSE = .84. Thus, for the present data, the assertion-negation bias was essentially due to a less favourable evaluation of positive-valence negations (as compared to positive-valence assertions), whereas negative-valence assertions were evaluated equally unfavourable as negative-valence negations.

More relevant for the present research, a marginally significant interaction of statement valence and onset of the first prompt was observed, $F(1,123) = 3.00$, $p < .09$, MSE = .84. See Chapter 3 for cell means and discussion.

Finally, an unexpected three-way interaction of cue valence, statement valence, and item list was observed. It was due to a pattern of agreement judgements such that, for *negations*, positive (as compared to negative) cues seemed to tend to further increase agreement with statements of positive valence, $p < .11$ for the pairwise comparison, but seemed to tend to further decrease agreement with statements of negative valence, $p < .11$ for the pairwise comparison; whereas for *assertions*, the respective cue effects did not approach significance, ps > .20 for the pairwise comparisons. For the three-way interaction, $F(1,123) = 4.30$, $p < .05$, MSE = .68. - No other effects occurred, Fs < 2.8, ps > .10.

To sum up, participants' dichotomous judgements were essentially determined by statement valence. Further, an assertion-negation bias was

observed. Whereas some evidence for the effectiveness of the onset-of-first-prompt manipulation was found, ordinal position of response, as well as cue valence, did not have much of an impact on judgements.

C2. Measure of both valence and speed (Study 2)

Next, participants' scores on a measure reflecting simultaneously both valence and speed of responding were analysed. Scores on this measure ranged from -2000 (indicating quick disagreement) to 2000 (indicating quick agreement); less extreme scores reflected less quick responding. With participants' eight scores each on the new measure as dependent variables, I conducted the same mixed-model ANOVA as before.

The analysis using the new measure revealed the same main effects that had been found in the dichotomous judgement data. Specifically, statement valence impacted scores such that more negative scores were found for negative-valence statements (M = -1169) than for positive-valence statements (M = 794), $F(1,123) = 544.65$, $p < .001$, MSE = 1791050. Scores for negations were more negative (M = -368) than scores for assertions (M = -7), $F(1,123) = 23.25$, $p < .001$, MSE = 1419578. Finally, scores were less negative when the first prompt had been presented very early (M = -109) than when it had been presented not very early (M = -267), $F(1,123) = 4.45$, $p < .04$, MSE = 1419578.

Further, the interaction of item list and statement valence found in the dichotomous judgements was confirmed, $F(1,123) = 19.02$, $p < .001$, MSE = 1791050, as was the three-way interaction of cue valence, statement valence, and item list, $F(1,123) = 4.56$, $p < .04$, MSE = 1407398. For each of

these interactions, the pattern of means led to the same conclusions as before.

Of greater interest for the present research, the previously marginally significant interaction of statement valence and onset of the first prompt turned out to be significant when using the new measure, $F(1,123) = 6.09$, $p < .02$, $MSE = 1791050$. In addition, a previously insignificant interaction of statement valence and ordinal position of response was found whereby for *negative-valence statements*, first responses were associated with less negative scores ($M = -1095$) than second responses ($M = -1243$); for the pairwise comparison, $p < .01$. For *positive-valence statements*, in contrast, no difference was found between first responses ($M = 760$) and second responses ($M = 828$); for the pairwise comparison, $p > .20$. For the interaction, $F(1,123) = 6.58$, $p < .02$, $MSE = 447434$. Thus, whereas the analysis of the dichotomous judgement data had not shown effects of ordinal position of response, such an effect was found with the new measure that took both valence and speed of responding simultaneously into account.

Most importantly, however, the analysis revealed a three-way interaction of statement valence, onset of first prompt, and ordinal position of response, $F(1,123) = 8.89$, $p < .004$, $MSE = 447434$. It was not qualified by higher-order interactions involving cue valence, Fs < 1.4, *n.s.* See Chapter 3 for a breakdown and discussion of this interaction.

Finally, two further interactions reached marginal levels of significance: firstly, the three-way interaction of item list, cue valence, and number of response, $F(1,123) = 2.77$, $p < .10$, $MSE = 317140$, and secondly, the three-way interaction of statement valence, cue valence, and number of response, $F(1,123) = 3.23$, $p < .08$, $MSE = 419914$. As these effects were

comparatively weak, they do not seem to warrant an in-depth discussion. - No other effects occurred, *p*s > .10.

To sum up, the measure that took speed as well as valence of responding simultaneously into account confirmed the effects found in the dichotomous judgement data. In addition, a significant three-way interaction of statement valence, onset of first prompt, and ordinal position of response was found that had not been observed for dichotomous judgements. The interaction was not affected by cue valence.