Congruency versus strategic effects in multimodal affective picture categorization

Paul M. C. Lemmens

ISBN–10 90–9019888–1
ISBN–13 978–90–9019888–0

Cover design and artwork by Marieke van der Linden, inspired by the cognitive science logo at `www.ai.ru.nl`.

Printed by PrintPartners Ipskamp, Enschede, The Netherlands.

Congruency versus strategic effects in multimodal affective picture categorization

een wetenschappelijke proeve op het gebied van de Sociale Wetenschappen

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de Rector Magnificus prof. dr. C.W.P.M. Blom,
volgens besluit van het College van Decanen
in het openbaar te verdedigen op maandag 31 oktober 2005
des namiddags om 3.30 uur precies

door

Paul Marcel Carl Lemmens
geboren op 1 oktober 1975 te Heerlen

Promotor:

Prof. Dr. G. P. van Galen

Copromotores:

Dr. A. de Haan
Dr. R. G. J. Meulenbroek

Manuscriptcommissie:

Prof. Dr. H. Bekkering (voorzitter)
Prof. Dr. K. R. Ridderinkhof (Universiteit van Amsterdam)
Dr. R. van Egmond (Technische Universiteit Delft)

Contents

CHAPTER 1

Introduction & overview

Our daily environment constantly claims our attention. When driving a car, for example, we continuously process all kinds of information on the basis of which we have to make intelligent decisions. Using as much information as possible in those situations has clear advantages, because the combined information from several modalities often produces an improved percept compared to information from one modality only (Calvert, Brammer, & Iversen, 1998). In cognitive science several views on the nature of human information processing exist. For example, with respect to the important question whether human information processing is either an active or a passive process, opposing views have been developed. According to some scientists, people actively search their immediate surroundings to select only those sources of information that are relevant for task performance (selection for action; Allport, 1987, also see Hommel, Müsseler, Aschersleben, & Prinz, 2001). According to others, the initial stages of information processing are purely passive and only after all the information impinging on our senses is taken in, the task–relevant components are being selected (selective attention; e.g. see Donk & Theeuwes, 2003; Pashler, 1998b; Theeuwes, 1994; Theeuwes, Godijn, & Pratt, 2004; Theeuwes, Kramer, Hahn, & Irwin, 1998).

Only very seldom will the environment stimulate only one of our sensory organs (Meredith, 2002). Usually next to visual stimulation — for most people the dominant modality (Ernst & Bülthoff, 2004; Stein & Meredith, 1993)— our hearing will also be active during information processing, resulting in bimodal stimulation of our cognitive system. However, often we also rely on, for instance, tactile feedback (a sticky keyboard) or even olfactory information to guide our multimodal interpretation of the surroundings in which we operate. Olfactory sensation, however, is mostly not used as a descriptor per se (“Smell the salty sea breeze”) but rather provokes an emotional or motivational context that possibly influences other, more rationally determined decisions (“Take a shower before you come by to discuss your paper’s grade”).

It may be even the case that a great paper does not receive proper recognition because of negative associations a reviewer has with smelly people. Such incompatibilities between the rational and emotional evaluation of situations are in abundance in daily life. These dimensions of decision making have been labeled “hot” and “cold” cognition, respectively (Abelson, 1963; also see Hoffman, 1986; Hollnagel, 1999; Sorrentino & Higgins, 1986a). Alternatively, consider, for example, driving your car in a hurry and approaching an intersection with traffic lights. During your approach you might hear the siren of an ambulance. If you have a red light when you reach the intersection, the interpretation of auditory and visual information are compatible: by law the red traffic light tells you to stop and the law tells you to give way to an ambulance driving with siren and horn. However, when the light is green, allowing you to reach your destination quicker, the light and the siren provide conflicting, incompatible information, usually resulting in an increased response–time latency to execute the appropriate response (here, pressing the brake). Now imagine being rescued successfully by an ambulance at some time in your past. This will likely change your emotion hearing the siren. Is it in this case easier for you (i.e., can you be quicker) to press the brake paddle of your car, or does your personal affective–evaluation of the siren not affect that response? In this example, the siren of the ambulance provided additional (new) information that required a different interpretation of a situation that you were confronted with while driving your car.

Usually the various sources of information available to people, not all bear relevance to their task performance. Indeed, in a lot of circumstances (multimodal) information presents itself in which part of that information is not relevant for a correct or appropriate execution of the task at hands. For instance, working with a computer often requires the use of a hierarchical menu–structure that can be very deep. An interface designer might consider adding sound to the computer interface to help users navigate within the menu structure; this sound, however, is not a necessity for proper navigation. The designer can use the pitch of a tone because pitch is a suitable auditory dimension to represent height (Walker, 2002; Walker & Ehrenstein, 2000). As a menu tree resembles a real tree, the designer might associate low(er) tones with menu items close to the roots of the tree. Going up higher into the branches and the leaves, the designer may use higher–pitched sounds. Surprisingly, research has shown that most users prefer the inverse relation between pitch and menu–depth: The deeper into the menu structure, the lower the sound should be (Walker & Ehrenstein, 2000). Thus, the original idea of high–pitched tones for items deep in the menu structure can be called an incompatible relation between pitch and position in the menu structure (at least for most people; B. Walker, personal communication, September 29, 2004; also see Walker, 2002), resulting in worse performance (measured in terms of errors committed or perhaps time to reach a certain position in the structure) compared to the compatible combination with lower sounds representing deeper positions in the menu structure.

The abovementioned examples of combined audio–visual information processing present in human task–performance, or processing of in general multimodal or multisensory stimuli form a normal part of our daily information–processing life. Meredith (2002) describes a multisensory stimulus as "an event which generates several independent physical 'energies' each of which is simultaneously detectable by different types of sensory receptors" (p. 32). Only because the human nervous system has been designed in an environment that generates those multisensory events, has it been possible that the physical energies produced or reflected by the events indeed do not go undetected. Meredith calls these physical energies the stimuli. The cross–modal aspect of such multimodal stimulation usually refers to the fact that the processes encoding the initially separate information–streams may affect each other in either beneficial or counterproductive ways, or stated in other words, that by–products of the processing of one sensory organ may carry over to the processing or end product of another sensory modality. The end result of a multimodal stimulus is therefore often different from a simple accumulation of the perception of each individual component of the stimulus (Meredith, 2002).

1.1 Cross–modal perception

Human sensory–modalities are selectively tuned to process certain stimulus characteristics. Gaver (1989) noted that auditory stimuli seemed specifically appropriate to relay information that changes (rapidly) and does not require direct, focused attention: Sound exists in time and over space. The visual system, on the other hand, is not equipped to handle rapidly changing information (certainly not in comparison with audition, the latter being able to distinguish events presented only 2 ms apart; vision, however, requiring at least 100 ms intervals between two stimuli; Fraisse, 1981) but vision can be used to describe static objects that require close inspection: Vision exists over time and in space. Although vision might be our dominant sense (Stein & Meredith, 1993), we rely on hearing as our only panoramic, long–range sensory system (King, Schnupp, & Doubell, 2001). In general, an auditory stimulus is processed around 30–50 ms faster than a visual stimulus (Fujisaki, Shimojo, Kashino, & Nishida, 2004; Stein & Meredith, 1993).

The modality–appropriateness hypothesis formulated by Welch, DuttonHurt, and Warren (1986) was the foundation for Gaver's observations. This hypothesis suggests that each sensory modality is capable of a variety of tasks but that they all have some functions that they can perform better than the other modalities. For instance, in a localization task with visual–auditory stimuli, vision will be biased over audition because vision is more appropriately designed to handle localization (Calvert et al., 1998). Note that a "more appropriate design" does not *necessarily* imply enhanced precision (Welch et al., 1986, p. 660). Alternatively, vestibular activation can change the apparent location of a sound (Stein & Meredith, 1993).

However, dominance of one modality over another one goes further than only for specific multimodal stimuli. For some people information processing in general is more visually oriented, usually enabling them to focus on visual task–aspects more than people who's processing is dominated by audition. Visually dominated people are therefore better capable of ignoring auditory information in a task in which both the visual and the auditory modality are stimulated (Giard & Peronnet, 1999). Dominance of specific modalities can change throughout the entire period of cognitive development. Children, for instance, have been found to be dominated much more by audition whereas most adults are visually dominant, perhaps reflecting a gradual change, over the course of adolescence, from language learner (requiring much more auditory attention) to object perceiver (requiring a visually oriented system, Sloutsky & Napolitano, 2003). Giard and Peronnet (1999) noted that auditorily–dominated adults are often musically skilled.

Modality appropriateness (Welch et al., 1986) and modality dominance are phenomena due to the very nature of multimodal stimuli. A well known example of a cross–modal effect caused by different information in the individual (unimodal) parts of a multimodal stimulus, is the McGurk effect (McGurk & MacDonald, 1976). This effect consists of a final percept that is different from either the visual or the auditory part alone. If participants hear "baba" while they see a speaker articulating "gaga" their cognitive system combines these two information–streams into "dada". Another example is the ventriloquism effect (Howard & Templeton, 1966) referring to various on–line as well as off–line manifestations of cross–modal spatial interaction in auditory–visual sensory conflict–situations (Bertelson, 1999) that show perceptual fusion despite spatial discrepancy between the auditory and visual sources. Vroomen and de Gelder (2000) showed an enhancement of the perception of a target when a high tone, embedded in a series of low tones, was presented synchronously with the visual target in the sequence of distracters. The enhancement did not depend on factors like rhythm or order–based anticipation.

These findings all point towards a cognitive interface subject to an interaction of initially separate auditory and visual information processing that seems to take place early on in the information–processing chain (Bertelson, Vroomen, de Gelder, & Driver, 2000; Fort, Delpuech, Pernier, & Giard, 2002; Giard & Peronnet, 1999).

1.2 Stimulus–response compatibility

So a computer–interface designer, thinking that a menu–tree grows into the air like a normal tree thus associating going up with higher pitch, created an interface–layout most users find awkward or in another word, incompatible. Stimulus–response compatibility (SRC) has been a major theme in (fundamental) experimental psychology as well as in human–factors research (Proctor & Reeve, 1990) and refers to a (in)comparability of some feature of a stimulus and a response that influences information processing. When two stimulus features have a relationship that is not according to a strong population stereotype (Fitts & Seeger, 1953; Fitts & Deininger, 1954), we generally speak of incompatible stimulus–features. Stimulus–response relationships not necessarily need be either compatible or incompatible. They may not even exist. For instance, if participants are instructed to press a left button if they see a red dot and a right button if they see a green dot, then color is the relevant stimulus–feature that determines the response. Because stimuli can be coded by color and responses are expressed as locations, no comparability exists between the relevant stimulus feature and the response and there can be no (in)compatibility between stimulus and response.

Now consider that the red or green dots are presented to the left or right of a centrally presented point of fixation. The stimulus now carries two potentially important features: the task–relevant feature still is color which, as indicated above, is dissimilar with the locational feature of the response. However, the lateralized presentation introduces a task–irrelevant feature to the stimulus that does overlap with the response. This setup results in two possible situations. Remember that participants were instructed to press the left button for a red dot. If this dot is presented to the left of the point of fixation, the location of the stimulus matches the location of the response which is called stimulus–response compatible. If, on the other hand, the red dot is presented to the right of fixation, this would result in a stimulus–response incompatible trial. In this setup participants typically respond around 50 ms–60 ms slower to stimulus–response incompatible trials than to stimulus–response compatible trials, an effect that has become known as the spatial Simon effect (Simon, 1969, 1990; Simon & Rudell, 1967).

In general, the properties of a stimulus that determine the response to be executed, usually specified in the instruction to the participants, are referred to as the relevant stimulus–aspect(s) or relevant stimulus–feature(s) (or "the relevant stimulus", abbreviated as S_r). Task–irrelevant stimulus–properties (S_i) are all remaining aspects of a stimulus that are not directly relevant for determining the appropriate response. Sometimes they are called task–irrelevant properties to emphasize that the properties are irrelevant for proper task–execution. Nevertheless, they may have an impact on the processing of the relevant aspects as exemplified in the spatial Simon effect. The relevant response–feature, or the feature distinguishing the responses, or just "the response" is the response feature (abbreviated as R) that allows the different responses to be distinguished. For instance, if one response button is located on the left of a response box and another one is located on the right of that box, R is location.

The relations between the relevant stimulus features and the response but also those between the relevant and irrelevant stimulus features are called either compatible or incompatible, or also congruent versus incongruent.[1] These relations can be phrased in terms of dimensional overlap between stimulus and response (Kornblum, 1994) where "dimensional" refers to the shared aspects of stimulus (S_i or S_r) and response (R), and overlap refers to the degree of similarity. For instance in the spatial Simon task there is dimensional overlap based on location of the irrelevant stimulus–feature and the response. In the type of research described in this thesis, similarity usually is all or none: shared stimulus and response features are very alike, or on opposite sides of the scale of similarity. Of course, dimensional overlap can exist between multiple features of a single stimulus.

---

[1] Within the SRC research arena, the distinction between compatibility, congruency, and consistency is usually upheld more stringent (Kornblum & Lee, 1995; Kornblum, Stevens, & Requin, 1999; also see Hommel, 1997, Footnote 1).

Note that although the term "dimensional overlap" was coined by Kornblum, Hasbroucq, and Osman (1990) as an integral part of their dimensional–overlap model of SRC phenomena, it is often used without direct reference to that model, both in this thesis as well as in the literature on this subject, solely referring to the similarity between S_r, S_i, and R. Also note that it may seem as if the relevant and irrelevant features of a stimulus are separate entities, certainly in the context of the multimodal stimuli where one modality is often used to present the task–irrelevant information and another modality exclusively presents the task–relevant features. This is, however, not necessarily always the case: The initially separate information–streams are assumed to be perceived as an integrated whole.

1.3 Affective information processing

Another topic briefly mentioned above was the impact of motivational context or affective appreciation on information processing. The processing of emotionally–labeled information, evaluative information, or affectively–valenced information (often called "hot" cognition, Abelson, 1963) was long considered to be completely separated from the cold rational decisions that humans make ("cold" cognition, Hollnagel, 1999), perhaps partly because of the computer metaphor (Shannon & Weaver, 1969) that dominated experimental psychology and human–factors research for a long time (Sorrentino & Higgins, 1986a). The labels affective valence, affective value, evaluative information, or affective property all refer to the same notion: some aspect, not otherwise specified, of a stimulus or a response carries a certain load that is processed more easily within an emotional ("hot"), rather than a rational ("cold") framework.

Common sense, however, seems to contradict this complete separation of hot and cold cognition: It is not just common–sense politeness to take a shower before an application meeting. Such a personal impression based on affect can easily influence decisions even if the application committee tries very hard not to be influenced by it. As another example, even for life–long partners a very sensible discussion on the spending of a minor amount of money can become a heated debate with a few words carrying an unexpected and often unintended emotional charge. In Western tonal music, the major/minor distinction is a powerful means to change or affect the (conscious and cold) perception of a piece from happy and positive to sad and negative (although exceptions exist; see Pittenger, 2003).

This well evolved (although occasionally volatile) domain of human–human interaction has often been used as starting point in human–factors reseach to improve the design of human–computer interaction. In a similar fashion as in experimental psychology (Sorrentino & Higgins, 1986b), the impact of emotion on human–human interaction has been ignored in human–factors research for a long period until Picard (1997) proposed the notion of affective computing by which she referred to all computing that displays and understands affect. Affective computing therefore seeks to improve human–computer interaction by incorporating the motivational and emotional factors that are so influential in human–human interaction.

1.4 Outline of this thesis

The studies that we report in this thesis are ordered chronologically, reflecting the transition of this project from a more or less purely human–factors research project (e.g. see Lemmens, Bussemakers, & de Haan, 2000, 2001; Lemmens, de Haan, & van Galen, 2003), to a more hybrid approach of basic experimental research with potential human–factors applications in mind. The critical evaluation of the research approach that we have adopted that we present near the end of this thesis in Chapter 5 indicates that extending the lifetime of the project might have resulted in a second move taking it even further away from human–factors research.

The next chapter (Ch. 2) elaborates on the research topics that were briefly touched upon earlier in this chapter. Based on an overview of the theories and findings in the stimulus–response compatibility literature, we argue that SRC phenomena can be used in the domain of affective processing, and that SRC paradigms can be exploited as a simpler way to investigate the impact of affect on cold cognitive processing compared to the direct but cumbersome measures of responses of the autonomous nervous system (blood–volume pressure, galvanic skin–conductivity, etc.) that are currently used in affective–computing research. We also argue that the major/minor distinction in Western tonal music can be used to transform families of affect–neutral earcons (which are brief musical fragment used in some man–machine interfaces to facilitate, for instance, navigation; Blattner, Sumikawa, & Greenberg, 1989) into variations that carry an affective load.

Chapter 3 examines in more detail the transformation of earcons into major and minor variations in the context of affective computing. We investigate whether the affective valence of earcons in major or minor mode (associated with positive and negative valence, respectively) can overlap with the affective property of the responses in a picture–categorization task. If the earcons and the responses have dimensional overlap, different congruency conditions can be expected. For instance, a combination of an earcon in major mode and a positive response would result in an affectively–congruent trial whereas an earcon in major mode and a negative response would be affectively incongruent. Finding differences in the response–time latencies to congruent trials versus incongruent trials would indicate that the transformation was successful in creating affectively–charged earcons. Insights into the conditions under which multimodal affective–congruency effects occur, are important to formulate targeted advice how to maintain affective correspondence in affective human–computer interfaces.

Chapter 4 explores the affective–congruency effect in more detail than was undertaken in the study reported in Chapter 3. In the study reported in Chapter 4 we use a slightly different experimental setup to try and replicate our earlier findings. We discuss that the multimodal affective–congruency effect represents a comparison across modalities of affectively–charged information. We argue that the multimodal research–approach provides a new method, in addition to existing visually–oriented paradigms, to investigate affective–information processing. We show that judgemental tendencies (see Klauer & Stern, 1992; Wentura, 2000) play an important role in realizing the multimodal affective–congruency effect. We contend that the judgemental tendencies are a cognitive strategy to simplify the procedure to select the instructed response.

A third and final experimental study, which we report in Chapter 5, critically evaluates the experimental set–ups used in the experiments described in Chapters 3 and 4. We argue that the class homogeneity of the sets of animals and inanimate objects is an important factor in studies that employ the animate/inanimate distinction (e.g., Brousseau & Buchanan, 2004) and that differences in the homogeneity of categories of stimuli may introduce artifacts in the multimodal affective–congruency effect. Several explanations of the congruency effect are investigated and we explore the extent of the influence of the homogeneity differences in the realization of the congruency effect.

In the general discussion, Chapter 6 (“Summary and conclusions”), the experimental findings reported in this thesis are recapitulated and discussed in relation to affective computing as well as to the main theories concerning affective–compatibility effects that are reported in the literature. We not only show how the paradigms that we used might be exploited in future stimulus–response compatibility research aimed to improve our insights into affective–information processes but also indicate the various ways in which our research can be applied to the cognitive–ergonomic domain of affective computing. We end the chapter with a description of the contours of a tentative model of hot and cold cognition that encapsulates our findings and lines of reasoning.

Chapter 2

Stimulus–Response Compatibility and Affective Computing: A Review[†]

> I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind.
>
> *Lord Kelvin (1824–1907)*

Abstract

Affective computing, a human–factors effort to investigate the merits of emotions while people are working with human–computer interfaces, is gaining momentum. Measures to quantify affect (or its influences) range from EEG, to measurements of autonomic–nervous–system responses (e.g., heart rate, blood pressure), to less objective self–reports. Here we claim that simple response–time measurements may be a viable alternative to (indirectly) measure the effects of affect on performance by providing a review of experimental paradigms and associated models of human information processing. In particular, we focus on stimulus–response compatibility paradigms that have provided important insights for human–factors research, for instance regarding the important role of the spatial layout of interface design on the efficiency of human task performance, to show that these paradigms can also be applied to investigate the role of affect in human–computer interaction.

[†] This chapter is submitted as: Lemmens, P. M. C., de Haan, A., van Galen, G. P., & Meulenbroek, R. G. J. (2004). *Stimulus–Response Compatibility and Affective Computing: A Review.* Manuscript submitted to Theoretical Issues in Ergonomics Science.

2.1 Introduction

Traditionally cognitive science and experimental psychology have investigated information processing within individuals (e.g. see Donders, 1868/1969; Kornblum et al., 1990; Sternberg, 1969). Since recently, however, the research investigating interindividual information–processing is gaining momentum (e.g. see Frith & Wolpert, 2003; van Schie, Mars, Coles, & Bekkering, 2004). Interindividual communication and information processing has been one of the pillars of human–factors research to improve human–computer interaction by trying to make it more like human–human interaction, for instance by using natural–language dialogue. Although often ignored, an essential component of human–human interaction is emotion: A wealth of non–verbal information can be conveyed by gestures or prosody. Even without verbal communication, humans often feel an immediate like or dislike for a person they meet for the first time, or, even between life–long partners communication can ignite a fierce debate with only a few sentences carrying an unexpected and (usually) unintended emotional charge. The success of emoticons in human–computer–human communication (email, chatting and online–presence software) corroborates the important role of emotion in human–human and human–computer interaction.

Originally human–computer interaction (HCI) did not involve emotion at all; the computer was seen as a binary machine devoid of any possibilities to position its responses between the extremes of zero and one and users simply had to adapt to that situation (Brave & Nass, 2003). A similar perspective was taken in experimental psychology where emotion (hot cognition) and reason (cold cognition) were long taken to be completely separated (e.g., Sorrentino & Higgins, 1986a). Picard (1997), therefore, argued that perception and appropriate presentation of affect by a computer might facilitate interaction with users and improve user performance in, as well as user satisfaction of, such an interface. She proposed to investigate Affective Computing which she defined as 'all computing that relates to, arises from, or deliberately influences emotions'.

Perhaps one of the reasons why hot and cold cognition have been viewed as separate is that the former is usually associated with responses on the level of the autonomous nervous system (e.g. heart rate, blood–volume pressure, skin conductivity, Brave & Nass, 2003) whereas the latter is associated with higher cognitive mechanisms. Most experiments measuring (the influence of) affect on HCI have therefore tapped into those autonomous nervous systems (the galvanic skin–response, GSR, Healey, Picard, & Dabek, 1998; Ward & Marsden, 2003; blood–volume pressure, BVP, Ward & Marsden, 2003; electromyograms, EMG, Scheirer, Fernandez, Klein, & Picard, 2002; or squeeze pressure on computer mice, van Galen, Liesker, & de Haan, 2004; van de Ven, 2002). Although these measures all have the benefit of directly measuring an affect–associated (autonomous nervous system) response, their downside is the currently rather cumbersome devices required for carrying out the measurements that, from a user's perspective, probably neutralise potential benefits of the addition of affect in an interface.

In this review we therefore propose to exploit stimulus–response compatibility (SRC) paradigms to study the role of affect in HCI as a viable albeit indirect alternative to the direct measures of EEG, GSR, BVP, and EMG. We will do so by first providing a brief overview of response–time paradigms and associated models of human information processing. Then the dimensional–overlap model and a taxonomy of stimulus–response compatibility paradigms is presented and the well–known Simon task is isolated to demonstrate the power of SRC in the domain of affective computing research. We complete our review by showing that earcons (Blattner et al., 1989) in major or minor mode as known from Western tonal music, may be a possible means to communicate emotions or affect in HCI. To the best of our knowledge this is a role for earcons that has not received much attention yet.

2.1.1 Brief historic overview

Donders (1868/1969) formulated a way to investigate human information processing by demonstrating that some tasks take longer to initiate than others. He proposed a stage insertion/deletion method in which a relatively simple task could be made a bit more difficult by requiring participants to do an extra 'computation' before executing a response. By carefully manipulating this extra computation, Donders reasoned that the time needed by this computation could be deduced by subtracting the time to initiate the easier task from the time to start the more difficult task, much like comparing the latency to generate the outcome of the sum 3 + 4 with that of the sum 3 + 4 + 5. To implement this general method in more detail, Donders defined three characteristic tasks that differed in the processes needed to correctly fulfill the task requirements. Type *a* reactions/tasks were simple RTs reflecting only sensory and motor processes. The *b*–type reactions were choice RTs with two stimulus and two response alternatives (including discrimination and response choice next to the sensory and motor processes of the *a*–type task), whereas the type *c* reactions were similar to a go/no–go paradigm (with discrimination processes but no response choice). This way the duration of, for instance, the response–choice process could be estimated by subtracting response times of *c*–tasks from *b*–tasks. This method of pure stage insertion was the first processing–stages approach to choice RT (Sanders, 1998).

However, difficulties were found creating sufficiently different type *b* and type *c* tasks as it appeared that both types of tasks required response choice. To overcome the methodological as well as statistical criticisms on Donders' subtraction method, Sternberg (1969) defined the Additive Factors Method (AFM, sometimes also referred to as Additive Factors Logic, AFL) that assumed, like the subtraction method of Donders, that response–time latencies consist of the sum of the intervals that a sequentially ordered modular set of processing stages take to translate a stimulus into response and start executing this response (Sanders, 1998). Each of these stages takes care of processing a certain aspect of the stimulus or response, possibly influenced by an experimental manipulation. The benefit of the AFM over Donders' method were the statistical assumptions that Sternberg formulated, creating a solid statistical framework in which experimental results from factorial designs could be interpreted in only a single way. Interactions between two (or more) factors signified that all factors acted upon the same information processing stage whereas additive effects implied that each involved factor influenced a different stage. Thus, during the heydays of the AFM around seven different stages were proposed or discovered, amongst which the stages of Feature Extraction, Stimulus Identification, Response Selection, and Motor Programming received most attention and consensus (Gopher & Sanders, 1984; Sanders, 1980, 1983, 1990, 1998).

However, these statistical premises, including ones derived from the basic assumptions, were violated quite easily. For instance, the stage–robustness criterion (Gopher & Sanders, 1984), a derived assumption arguing for constancy of the stage structure over many variations (including number) of factors over experiments, was violated using stimuli typically used in experiments employing stimulus–response compatibility paradigms (Ridderinkhof, van der Molen, & Bashore, 1995; Sanders, 1998). Dual–route architectures were proposed (de Jong, Liang, & Lauber, 1994; Kornblum et al., 1990; Ridderinkhof et al., 1995) to overcome the limitations that the AFM imposed on SRC experiments.

2.1.2 Stimulus–response compatibility

A study by Fitts and Deininger (1954) is one of the most often cited studies to introduce stimulus–response compatibility research. They instructed one group of participants to move a stylus to the left if, for instance, a stimulus light on the 9 o'clock position of circle of lights was turned on whereas another group of participants was instructed to move the stylus to the right in response to the same stimulus. It is clear that the task was more difficult for the second group than for the first group of participants because of the spatial incompatibility between the relative position of the stimulus versus the direction of the response. Although Fitts and Deininger termed the condition of the second group S–R mirrored, these types of stimulus–response assignments have become known as incompatible stimulus–response combinations.

The widely used Eriksen flanker task (Eriksen & Eriksen, 1974) is another example of a task that capitalizes on stimulus–response compatibility effects. In a variation with arrows, the flanking arrows could point in the same or opposite direction as the centrally presented target arrow, resulting in compatible or incompatible trial types,[1] respectively (Ridderinkhof et al., 1995). Compatibility of the direction of the target arrow and the flankers influenced response–time latencies: RTs to incompatible trials were significantly slower than those to compatible trials. The spatial Simon task (Simon, 1990) is also considered to elicit SRC effects, as is the Stroop task (Stroop, 1935; Proctor & Reeve, 1990). These tasks will be discussed below.

Stimulus–response compatibility has been a major topic in human–factors research since the work of Fitts and Seeger (1953; also see Fitts & Deininger, 1954). Laxar and Olson (1978), for instance, showed that SRC is an important factor for judgements on computer–generated diagrams of submarine movements. Burke, Gilson, and Jagacinsky (1980) highlighted the importance of SRC in multimodal displays intended to relieve visual workload. Brown (1988) wrote an entire chapter on SRC in a book on HCI–design guidelines and Proctor and Reeve (1990) already

---

[1] In contrast to some researchers who explicitly use different terms for different the different relations between S_r, S_i, and R like congruency, consistency, correspondence as well as compatibility, for the present review we will use the words interchangeably.

dedicated a part of their book on the theme of SRC and human–factors. Like Andre and Wickens (1990), Kantowitz and Campbell (1996) stressed the impact of stimulus–response compatibility on a pilot's work load. For directional compatibility, Worringham and Beringer (1998) showed that visual–field compatible conditions were associated with shorter response time latencies, movement and homing–in times. Summarising, Vu and Proctor (2003) stated that spatial SRC effects are robust (see Hommel, 1997) and occur across different sensory modalities, stimulus modes, response modalities, and display arrangements. Thus, stimulus–response compatibility effects are widespread throughout cognitive performance and in human–factors research the general rule to maintain (spatial) correspondence emerged as an important interface design guideline.

Kornblum et al. (1990) tried to systematise the different paradigms employed in stimulus–response compatibility research by classifying the stimulus–response relations (see Fig. 2.1) of each paradigm into a taxonomy of so–called stimulus–response ensembles. To accomplish this, they formulated the notion of dimensional overlap (DO) referring to 'the degree to which sets of items are perceptually, structurally, or conceptually similar' and presupposed that 'dimensional overlap affects performance whether the overlapping dimensions are task–relevant or not' (Kornblum, 1994, p. 130). Although dimensional overlap did not exclude degrees of overlap on a continuous scale, Kornblum et al. (1990) varied the relations between S_i, S_r, and R in a binary fashion, resulting in eight possible ensembles.[2]

For instance, a task having a dimensional overlap (DO) between S_r and R is called a Type–2 ensemble (see Table 2.1). A typical Type–2 ensemble is the spatial compatibility study by Fitts and Deininger (1954; also see Fitts & Seeger, 1953) that we described above. In these types of S–R ensembles the relevant stimulus feature S_r is location, which is also the feature that distinguishes the individual responses R. The Stroop task (Stroop, 1935) belongs to the most complex ensemble of the taxonomy: the irrelevant colour–words create have an overlap with the relevant ink–colour ($S_i - S_r$) as well as with the response (pronounce the name of the ink–colour; $S_i - R$). Because there is also overlap for $S_r - R$ (the ink–colour and its name), the Type–8 ensemble, to which the Stroop task belongs, is characterised by

---

[2] Recently Stevens and Kornblum (2000) extended the number of possible ensembles to around twelve to also include response–effect compatibilities (e.g. Beckers, De Houwer, & Eelen, 2002; Koch & Kunde, 2002; Kunde, Hoffmann, & Zellmann, 2002; Kunde, 2003; also see Kornblum & Stevens, 2002).

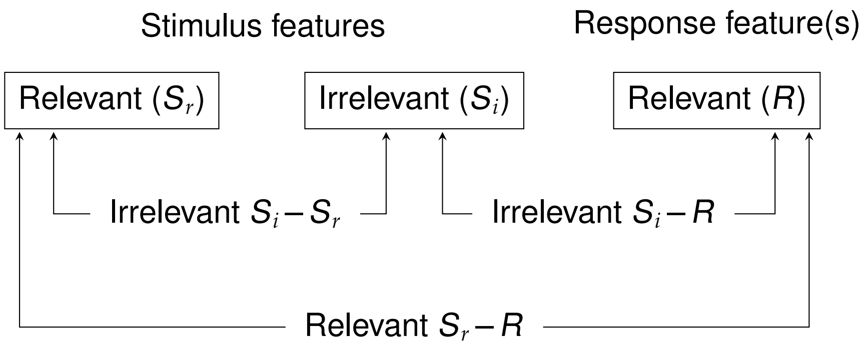

Figure 2.1: Three possible relations between the relevant (S_r) and irrelevant (S_i) aspects of the stimulus and the relevant response feature (R) used to create eight ensembles in the dimensional overlap taxonomy (Kornblum et al., 1990). Figure adapted from Fig. 2 in Hasbroucq and Guiard (1991).

dimensional overlap between all combinations of relevant and irrelevant stimulus features and response.

Compared to the original spatial–compatibility study carried out by Fitts and Seeger (1953), the spatial Simon effect (Simon, 1969; Simon & Craft, 1971; Simon, Hinrichs, & Craft, 1970; Simon & Rudell, 1967; also see Simon, 1990) is only slightly different: whereas Fitts and Seeger studied location as the relevant stimulus feature, for the spatial Simon effect location is task–irrelevant. In Kornblum et al.'s (1990) taxonomy this conforms to a Type–3 ensemble with no overlaps between S_r and R or between S_i and S_r. However, for $S_i - R$ there is overlap because the irrelevant stimulus feature is location and location also distinguishes between the possible responses.

Accompanying the taxonomy, Kornblum et al. (1990) introduced a dual–route processing model of SRC effects (see Fig. 2.2) consisting of a stimulus–identification stage and a response–production stage (where these stages indeed have some characteristics from the stages in the AFM, cf. Kornblum & Lee, 1995), the latter stage mediating the effects of Type–3 ensembles like those found in the spatial Simon task. This stage consists of two branches in which one branch is responsible for an automatic activation of a response based on the dimensional overlap between S_i and R.

Table 2.1: A taxonomy of S–R ensembles, illustrating the various combinations of dimensional overlap with relevant and/or irrelevant dimensions of the stimulus and/or the response. Table adapted from Table 1 in Kornblum (1994).

Ensemble type	Overlapping ensemble dimensions		
	S–R dimensions		S–S dimensions
	Relevant	Irrelevant	
1	No	No	No
2	Yes	No	No
3	No	Yes	No
4	No	No	Yes
5	Yes	Yes	No
6	Yes	No	Yes
7	No	Yes	Yes
8	Yes	Yes	Yes

However, before the actual response can be executed, the *correct* response must be identified; this response identification takes place in the other branch using, for instance, an identity rule. If the automatically activated response is different from the correct response, the former is aborted and the program for executing the correct response is retrieved and executed (Kornblum, 1994). The process of aborting an incorrect automatic response results in the observed response–time latency differences for a compatible trial compared to an incompatible trial.

The stimulus–identification stage was constructed in a similar fashion to the response–selection stage with two separate routes converging to create a stimulus vector (Kornblum, 1994). The stimulus–identification stage mediates SRC effects, for instance, for Type–4 ensembles. This model was (Kornblum, 1994; Kornblum et al., 1990; Kornblum & Lee, 1995) and still is (Kornblum et al., 1999; Kornblum & Stevens, 2002; Stevens & Kornblum, 2000) successful in explaining many of the SRC phenomena.

To summarise, so far we have discussed the chain of stages in information–processing that are associated with the AFM, the successor to Donders' stage–insertion and –deletion method. The assumptions of the AFM failed, however, for many of the types of tasks and stimuli that

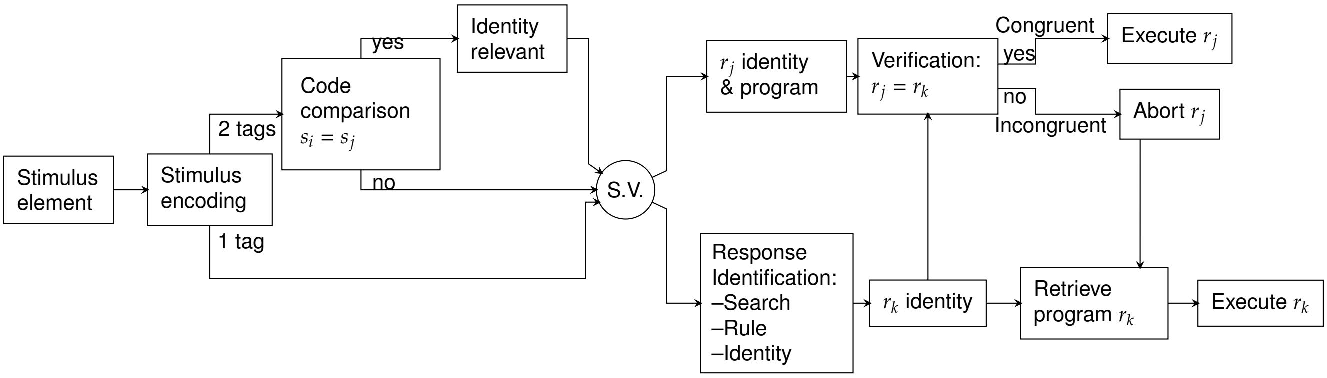

Figure 2.2: The dimensional overlap processing model. The stimulus vector (S.V.) marks the cutpoint in the network. To the left of the cutpoint is the stimulus identification stage; to the right the response production stage. s_i = a stimulus attribute that overlaps with another stimulus attribute; s_j = a stimulus attribute that overlaps with a response attribute; r_j = an automatically activated response; r_k = the correct response. Reproduced, with permission, from Figure 1 (including caption) in Kornblum and Lee (1995).

are often used in SRC studies and various notions of dual–route models succeeded it. Out of these models we presented the dual–route model of Kornblum et al. (1990) that is associated with their taxonomy of stimulus–response compatibility ensembles. We showed that SRC has been an important factor in many human–factors studies. We now discuss affective–information processing and a variation of the Simon task to show that SRC paradigms can be used in the domain of affective computing.

2.2 Affective information processing

Earlier in the history of experimental psychology, research on hot cognition (Abelson, 1963) was, to a large extent, separated from that on cold cognition, perhaps due to the influence of the computer metaphor (Shannon & Weaver, 1969). Affect, the common denominator of emotion, mood, and sentiment, was seen as a byproduct or end product of cognitive appraisal (Schacter and Singer, 1962, cited in Hoffman, 1986). During the 1980s the balance returned to a more favourable position for affect research: the complexity of the interaction of affect and cognition became appreciated. Zajonc (1980) argued for independent operation of the affective and cognitive systems, whereas Piaget (1981) related the amount of intellectual effort and the rate of knowledge acquisition to approach and avoidance situations. More recently the computer metaphor and hot cognition have been brought closer with the advent of affective computing (Picard, 1997), which is one of the main motivations of the present review.

Kuhl (Sorrentino & Higgins, 1986a, Ch. 14) distinguished affect or emotion on the one hand and cognition on the other by emphasising the representative relation of cognitive processes with real–world objects and facts, whereas affective/emotional processes evaluate the personal significance of those objects and facts. Brave and Nass (2003) also underlined the object–directedness of emotions, citing Frijda (1994): 'Emotions are intentional. They imply and involve relationships with a particular object'. Moods, on the other hand, are non–intentional, not object directed and are experienced as more diffuse, global, and general. Emotion is also different from moods on a functional level: as a reaction to a stimulus, emotion can bias a certain action for a short period of time. Moods, in contrast, bias strategies and processing over longer periods periods of time (Brave & Nass, 2003). Note that the influence of affect as a reaction to a relevant stimulus does not need to result from that stimulus per se: Hoffman (1986) proposed a set of processes, including semantic interpretation, as preparatory transformations that can put a stimulus in a form that allows affect to operate.

Affect can influence subsequent information processing by initiating it, terminating it, accelerating it, or disrupting it (Hoffman, 1986); it can influence global performance (e.g. positive affect facilitates cognitive flexibility and efficiency in thinking and problem solving, Brave & Nass, 2003), but is also effective in capturing attention, improving memory, recall, and category accessibility (Hoffman, 1986), and it can result in mood–congruent and affect–congruent decision making (Brave & Nass, 2003; Klauer & Stern, 1992). Affect can be caused by moods and sentiments, by previous emotional states, by specific goals or more abstract needs (Brave & Nass, 2003), or by judgemental tendencies (Klauer & Stern, 1992; Wentura, 2000).

2.2.1 Affective compatibility

Given the many global influences that affect can have on cognitive information processing, it is not surprising that in the last decade an increasing number of researchers have been putting much effort in understanding the cognitive underpinnings of affect. One way to carry out this type of research is to establish influences of affect on cold cognition using the conflict and interference tasks used in SRC research. If the two systems are separate, non–connecting mechanisms, then no effects of either conflicting or non–conflicting information in the affective domain should be observed in cold cognitive processing. On the other hand, affective congruency effects (i.e. faster processing of a stimulus–response mapping that has equivalent affective valence either within the stimulus or between the stimulus and the response) or incongruency effects would indicate an interface between hot and cold cognition.

One of the earliest attempts to investigate this issue was by Fazio, Sanbonmatsu, Powell, and Kardes (1986) who created a priming task in which the relation between prime and target was not based on semantics or phonetics (e.g. see Neely, 1991) but was based on affective valence. For instance, the prime–target pair SUMMER–HONEST was considered to be affectively congruent whereas CANCER–HONEST was considered affectively incongruent. This affective–congruency effect actually is an affective–priming effect consisting of the difference in mean response–time latencies to the affectively incongruent and congruent prime–target pairs. Fazio et al. (1986) indeed observed an affective priming effect.

This finding has been replicated in many other studies (De Houwer, Hermans, Rothermund, & Wentura, 2002; De Houwer & Randell, 2004; Hermans, De Houwer, & Eelen, 2001; Klauer & Musch, 2001; Moors & De Houwer, 2001; Wentura, 2000; but see Klauer & Musch, 2001, and Spruyt, Hermans, Pandelaere, De Houwer, & Eelen, 2004, for some exceptions). Affective congruency effects in a Stroop task were investigated by Rothermund and Wentura (1998). De Houwer and Eelen (1998) extended the spatial Simon task to the domain of affective information processing (see De Houwer, 2003; De Houwer, Crombez, Baeyens, & Hermans, 2001; De Houwer & Eelen, 1998; De Houwer, Hermans, & Eelen, 1998; Voß, Rothermund, & Wentura, 2003). Beckers et al. (2002) used the perception and action framework from the theory of event–coding (Hommel et al., 2001) to show an affective response–effect using trained adverse–response effects to certain responses that in a test phase were compatible or incompatible with the affective valence of the stimuli. Lemmens, de Haan, van Galen, and Meulenbroek (2004b) employed multimodal stimuli carrying an affective load obtaining a multimodal affective congruency effect in several studies (e.g. see Van Esch–Bussemakers, 2001; Lemmens et al., 2001; Lemmens, de Haan, & van Galen, 2004).

In a series of categorisation and naming studies, De Houwer and colleagues (De Houwer et al., 2002; De Houwer & Randell, 2004) showed that the semantic system seemed to be involved in the processing of affective information. They proposed a distributed model of semantic memory in which affective valence as well as semantic information are encoded as distributed patterns of activation across a set of processing units.

However, formally many of the affective priming studies cannot be considered stimulus–response compatibility effects because the temporal order of the (sequential) presentation of prime and target is different from that of relevant and irrelevant stimuli (that are presented simultaneously). The usefulness of SRC for affective computing would be better shown using a task that does fit under the umbrella of SRC. Such a task is the affective Simon task.

2.2.2 The affective Simon task

In the DO taxonomy of Kornblum et al. (1990), the Simon task is classified as a Type–3 ensemble that is characterised by an overlap between the irrelevant stimulus and the response ($S_i - R$) but no overlap between the relevant stimulus and response ($S_r - R$) or the relevant stimulus and the irrelevant stimulus ($S_i - S_r$). The $S_i - R$ overlap in Type–3 ensembles is responsible for the increase in response times when S_i is incompatible with R. For instance, in a spatial Simon task the relevant stimulus feature is colour that, by instruction, should be translated in lateralized responses (no overlap possible between colour and location). However, the stimuli are presented on the left or right of a fixation point, introducing a task–irrelevant location feature in the stimulus that has a dimensional overlap with the location feature of the responses; within the stimulus, however, there still is no overlap of S_r and S_i.

De Houwer (1998) and De Houwer and Eelen (1998) observed that the generalised feature–unspecific setup of overlaps for Type–3 ensembles did not limit the Simon effect to the spatial domain (also see Kornblum & Stevens, 2002). De Houwer (1998) used the Simon paradigm to investigate task–irrelevant semantic overlap in $S_i - R$ whereas De Houwer and Eelen (1998) used the paradigm to investigate the impact of affective valence on information processing (also see De Houwer et al., 2001).

The affective Simon task (De Houwer et al., 2001; De Houwer & Eelen, 1998) was constructed in a similar fashion as the regular spatial Simon task. Instead of a spatial or semantic feature (De Houwer, 1998), in this case it was affective valence that created an overlap between S_i and R.

Participants were instructed to pronounce the word POSITIVE when presented with a noun and to pronounce NEGATIVE when seeing an adjective. In this case the feature that distinguished the possible responses was affective valence and the relevant stimulus was grammatical category. The dimensional overlap of $S_i - R$ was created by using words in either grammatical category having established affective valences (Hermans & De Houwer, 1994), for instance, FRIEND and HONEST or THIEF and STUPID. Like the spatial Simon effect, this setup resulted in two different types of stimuli. For instance, the stimulus FRIEND has a congruent $S_i - R$ relation: FRIEND has a task–irrelevant positive affective valence overlapping with the relevant response feature also having a positive affective valence. On the other hand, the stimulus THIEF was considered to be incongruent, because THIEF has a negative affective valence whereas the appropriate response has positive affective valence.

This way, De Houwer and Eelen (1998) created a paradigm fulfilling all three criteria of the Type–3 ensembles characterizing the generalized Simon paradigm: (a) no $S_r - R$ overlap (grammatical category vs. affective valence), (b) no $S_i - S_r$ overlap (affective valence vs. grammatical category), and (c) the overlap of affective valence of the stimulus words and the affective property of the response could result in compatible or incompatible relations between S_i and R. Comparing affectively congruent trials to affectively incongruent trials showed a congruency effect of around 50 ms (De Houwer & Eelen, 1998, Exp. 2).

Extending the paradigm, De Houwer et al. (2001) and Tipples (2001) found affective–congruency effects for pictorial stimuli using different types of relevant stimulus attributes (grammatical category and semantic category; also see De Houwer et al., 2001) with response words that only indirectly carried a positive or negative valence (cf. FLOWER and CANCER vs. POSITIVE and NEGATIVE, De Houwer et al., 2001, Experiment 3; Tipples, 2001). These findings established the affective Simon task as a fruitful and useful variant of the Simon paradigm to study the processing of affective information.

In our view, the setup just described makes the Simon task an ideal candidate for research in the field of human–factors because it is a well defined experimental task that mimics real–world circumstances closely. In real life, more often than not, stimuli carry multiple features of which only one is strictly relevant for the task to be executed. The other features, however, can affect task or response execution because they somehow share information with the response to be executed without users being aware of the influence. The Simon task fulfills these criteria because without explicit notification (which is necessary in Fitts and Seeger's, 1953, spatial compatibility task) one can manipulate the task–irrelevant (S_i) location or semantics. In order to study affective computing in a systematic manner one can use affect as S_i, in a task (reflecting a specific $S_r - R$ translation) that is affect–unrelated (i.e. at least for the participants).

Findings of spatial compatibility effects have resulted in interface–layout design guidelines to maintain spatial correspondence. In a similar fashion we propose to exploit affective Simon effects for affective–computing purposes by taking into account more general affective–congruency effects, for instance, when using affectively charged earcons in human–computer interfaces. In the following section we argue that the major/minor distinction in Western tonal music can be used to transform (families of) affect–neutral earcons to variants in major or minor mode to facilitate or support the presentation of emotionally loaded messages. These affectively loaded earcons can then be carefully employed in multimodal affective human–computer interfaces.

2.3 The major/minor distinction

Sounds can evoke powerful emotions. Consider, for instance, a loud bang that will startle us or the shrieking of fingernails on a blackboard that most people find highly unpleasant. Such sounds often trigger responses ranging from immediate escape to making every effort to stop the noise. In emotion research these types of sound are known to elicit generalised arousal effects, orientation (flight or fight) reactions, and related emotional responses that are referred to as basic emotions whereas cognitive emotions (e.g. indignation or desire) follow attributions, that is, higher cognitive processing (Van Egmond, Desmet, & Van Der Helm, 2004). Within the realm of music processing, an important distinction related to emotion is that of major mode versus minor mode.

In western tonal music the major mode is presumed to have a default association of a happy, merry charge, whereas the minor mode is mostly labeled as sad (Crowder, 1984; Gregory, Worrall, & Sarge, 1996) but also as dreamy and sentimental (Hevner, 1936). Crowder (1984) described three possible ideas for explaining why major and happy go together.

Because (a) the higher partials of major chords can be observed in nature more often than those of minor chords and the positive connotation is therefore derived from a greater 'naturalness' of a major chord. Alternatively, von Helmholtz (1885/1954) hypothesised that (b) major chords are preferred over minor chords as a special case of a preference of consonance over dissonance.[3] Finally, (c) the evaluative connotations can be a cultural convention that arose by accident. Which of these proposals is the right one still seems to be under debate (Crowder, Reznick, & Rosenkrantz, 1991, p. 188). However, the emotional connotation of major and minor is a stable convention that, contrary to initial believes (Heinlein, 1928), does survive decontextualization (Crowder, 1984) even for chords presented as briefly as 300 ms (Crowder, 1985a, 1985b).

Crowder (1985a) used artificially generated sine–wave major and minor triads to investigate the sharpness of the categorical distinction of major and minor. Creating a nine–step continuum that musically skilled and musically unskilled participants had to judge on perceived majorness, he showed large interindividual differences with an obvious relation between musical sophistication and the reliability of the categorical discrimination. By comparing Experiment 1 and Experiment 2, Crowder (1985a) showed that musical skill was merely reflected in the sharpness of the distinction (i.e. musically unskilled participants were equally proficient in assigning major and minor labels to the extremes of the continuum but showed greater uncertainty at the centre of the continuum). This finding was corroborated by Howard, Rosen, and Broad (1992) who also found a strong positive relationship between musicality and the steepness of the psychophysical function of each participant for the major–minor continuum. Musically skilled participants showed steeper slopes, reflecting a sharper distinction between major and minor, than musically unskilled participants. Crowder explained the interindividual differences mostly by referring to the artificial nature of the sounds, possibly making the discrimination hard for most of the subjects. Crowder (1985b) elaborated on these experiments by having participants judge on a happy/sad dichotomy next to a major/minor categorisation resulting in equivalent performance for the happy/sad instruction and the major/minor instruction. Kastner and Crowder (1990) showed that children only three years old recognized the conventional connotation to a statistically significant

---

[3] Related to this proposal was the interesting alternative formulation by Montani (1945) who claimed that the lowered middle third of a minor triad produced castration anxiety, mediating the negative connotation of minor chords (cited in Crowder, 1984).

degree by pointing to happy or grumpy cartoon faces after hearing a short melody in either major or minor mode. Infants of around 6 months of age, however, did not show the major/minor distinction when tested in a gaze–tracking study (Crowder et al., 1991).

In sum, even for musically untrained listeners the major/minor distinction has the conventional connotation of positive and negative charge, respectively (Crowder, 1984), even for chords presented very briefly and without any musical context. With care, these findings suggest that the positive and negative affective valence of major and minor chords can be used as a dichotomous variable in experiments investigating the influence of affect on performance. Lemmens, Bussemakers, and colleagues (e.g. see Van Esch–Bussemakers, 2001; Lemmens et al., 2000, 2001, 2004b) have used the major/minor distinction repeatedly to construct earcons that carried an affective charge, in various studies that involved a simple picture–categorisation task. They found that the positive affective valence of the major mode (expressed as, e.g. a C5 triad) resulted in congruent and incongruent dimensional overlap between an earcon in major mode and positive or negative categorisation responses. For instance, negative responses to a picture accompanied by an earcon in major mode (negative response incompatible with positive valence of major mode) were significantly slower responded to as compared to responses to the same picture if that picture was accompanied by a minor chord (negative response compatible with negative valence of minor mode). Lemmens et al. (2004b) therefore suggested that the major/minor distinction can be used a basic musical aspect (among others like pitch, timbre, or register; also see Blattner et al., 1989; Brewster, Wright, & Edwards, 1994; Pramana & Leung, 1999) to transform affect–neutral earcons into variants in the major mode (associated with positive affect) or in minor mode (associated with negative affect). They suggested that this transformation seems especially useful for affective–computing research.

2.4 Discussion

To recapitulate, we have shown that the method of pure stage insertion and deletion by Donders (1868/1969) and its successor the AFM (Sternberg, 1969) were successful in determining a sequential set of steps of human information processing. However, because the statistical premises of the AFM were violated in typical SRC experiments (Ridderinkhof et al., 1995; Sanders, 1998), the AFM was supplanted by dual–route models that were proposed, for instance, by Kornblum et al. (1990), de Jong et al. (1994), and Ridderinkhof (2002).

Kornblum et al. (1990) tried to systematise the many tasks, including flanker tasks, Simon tasks, and Stroop tasks, that fitted under the umbrella of stimulus–response compatibility (SRC) research (Proctor & Reeve, 1990) by proposing the notion of dimensional overlap and a taxonomy of eight types of stimulus–response ensembles. We showed that SRC has been an important theme in human–factors research primarily restricted to the spatial layout of interfaces (Andre & Wickens, 1990; Vu & Proctor, 2003). We argued that SRC research might form a beneficial addition to the direct but complex measures of affect that are employed in current affective–computing investigations (e.g., Healey et al., 1998; Scheirer et al., 2002; Ward & Marsden, 2003) and we contend that affective–compatibility effects are a viable alternative to investigate affective–information processing.

Findings of prolonged response times to incongruent prime–target pairs and in trials with incongruent $S_i - R$ relations in an affective Simon task all strongly suggest maintaining affective correspondence in HCI seems as important as maintaining spatial correspondence (also see Lemmens et al., 2004b). We discussed affective priming research that hinted a the involvement of semantic processing in the realisation of many affective compatibility effects.

We discussed that sound can evoke powerful emotions that often drive instinctive reactions: Loud bangs or shrieking noises provoke immediate avoidance behaviour. On the other hand, the sounds of explosions and missiles being fired are important aspects of computer gaming that are often beneficial for performance, so sound does not necessarily hinder performance (also see the arousal effect, e.g., Keuss, van der Zee, & van den Bree, 1990). So in some situations adding sound to an interface seems to enhance performance while it hinders performance on other occasions. It seems important, therefore, to maintain correspondence between the sound and the event that it accompanies. In some cases this seems trivial advice, for instance, the action of dropping a bomb in a computer game should be accompanied with a whistling sound with decreasing pitch. On the other hand, with pitch as a navigational aid in a tree of menu options, moving up towards the leaves (the actual menu options that result in, e.g. a pop–up window when selected) should be accompanied by decreasing pitch (Walker & Ehrenstein, 2000; Walker, 2002).

Our own recent research has shown that the major/minor distinction in Western tonal music is relatively easy way to investigate multimodal affective–information processing (Lemmens et al., 2001, 2004b) because the major mode commonly carries a positive emotional load whereas the minor mode is usually associated with a negative load (Crowder, 1984, 1985a). We propose that the major/minor distinction can be used to transform earcons, with care, into variations capable of presenting positive or negative emotion. In this way, stimuli or events in an interface that do not carry an emotional load can be enhanced with auditory accessory stimuli to add an affective component. For instance, the long lists of emoticons that can be displayed in many email or online–presence software could be enhanced with earcons in major and minor mode to more easily differentiate between emoticons representing positive or negative emotions, respectively.

We would like to stress, however, that information processing can be delayed by the addition of sound, even when correspondence is maintained, because in any display the addition of information (whether redundant or complementary) adds a component of distractibility to the display that can detract attention away from processing the most important information. For instance, it is known that in operating–theatres the alarms of equipment monitoring the health of the patient under surgery are often shut off because they distract the surgeon (Meredith, Edworthy, & Rose, 1999; also see Van Egmond, 2004).

Of course, given the young age of the fields of affective computing and research on affective–compatibility effects, more research is needed to improve our understanding of the cognitive representation of affect and how these insights into this domain can be exploited in interface design. Pending these future research–efforts, we propose, in parallel with the argument to maintain spatial compatibility in human–computer interface display designs (Andre & Wickens, 1990; Vu & Proctor, 2003), that maintaining affective compatibility might be an important first step to ensure optimal performance in acoustically supported interfaces using affect–loaded messages.

Chapter 3

Emotionally Charged Earcons Reveal Affective Congruency Effects[†]

The most exciting phrase to hear in science, the one that heralds new discoveries, is not *'Eureka!'* but rather *'hmm....that's funny...'*

Isaac Asimov

Abstract

The domain of cognitive ergonomics has clearly profited from research into stimulus–response compatibility (SRC) conducted by experimental psychologists. In particular, the finding that performance speed is facilitated under spatially compatible stimulus–response relationships has found its way into guidelines for how to design an appropriate lay–out of human–computer interfaces. The search for SRC effects has not been restricted to spatial dimensions but has also been extended to colour and sound. In particular, the affective valence of sounds seems relevant for affective–computing research aimed at investigating how emotions can be communicated effectively in human–computer interaction. In the present study we show in a picture–categorisation task that the affective connotation of earcons in major and minor mode (representing positive and negative valence, respectively) can be congruent or incongruent with other task aspects. Our findings show that it is important to carefully investigate human–computer interfaces for potential affective–congruency effects because these can either facilitate or inhibit user performance.

[†] This chapter is submitted as: Lemmens, P. M. C., de Haan, A., van Galen, G. P. , & Meulenbroek, R. G. J. (2004). *Emotionally Charged Earcons Reveal Affective Congruency Effects.* Manuscript submitted for publication to Ergonomics.

3.1 Introduction

Research into (stimulus–response) compatibility effects has been a major topic in human–factors research, for instance, to investigate the optimal display–arrangements for interface–layouts. Stimulus–response compatibility (SRC) effects refer to findings of improved task performance for stimulus–response mappings that are more natural (or compatible) compared to less natural (incompatible) mappings (Proctor & Reeve, 1990). For instance, Fitts and Deininger (1954) instructed one group of participants to move a stylus to the left if a light on the 9 o'clock position of a dial would light up, whereas another group of participants was instructed to move the stylus to the right. The latter group experienced a stimulus–response incompatible mapping and they showed significantly prolonged response–time latencies because of the spatial incompatibility between the relative position of the light and the direction of the response. Originally, much of the research into SRC phenomena involved relatively simple domains like location or colour (e.g. see Kornblum et al., 1990, for an overview), but effects of semantic compatibility have been found as well (e.g. the semantic Simon effect, De Houwer, 1998). Recently, many instances of affective–compatibility effects have been reported (De Houwer & Eelen, 1998; De Houwer et al., 2001, 2002). Such findings suggest that not only spatial correspondence is important in interface designs, but that other aspects of human–computer interfaces may also be subject to effects of stimulus–response compatibility. The affective–compatibility effects specifically might be relevant for affective–computing research.

Affective–computing research (Picard, 1997) investigates the merits of incorporating aspects of human emotion into human–computer interfaces to improve the quality of human–computer interaction. To do so, its research efforts are twofold (Hollnagel, 1999). On the one hand, interfaces must be developed that can recognise the current affective state of its users. These interfaces, on the other hand, should be complemented by interfaces that can (also) present affectively charged messages. Regarding the first type of interface, already several devices have been developed that are capable of measuring and exploiting the affective state of their users. Healey et al. (1998), for instance, created a device that chooses the music that it will play based on the mood that it infers from measurements of the responses of autonomous–nervous systems (e.g., heart rate or galvanic skin–conductivity). Research into interfaces that can present affect, however, seems less prominent. In our view this is unfortunate, because especially in this type of interface, where two independent sets of affect interact —a sender (the computer) versus a sender–perceiver (the user)— there is ample opportunity for affective incompatibilities to come about.

In our research on multimodal information processing we have been investigating the effects of adding redundant sound in human–computer interfaces. We investigated the effects of earcons (Blattner et al., 1989) that carry task–irrelevant, redundant information on, for instance, reaction–time latencies to a task that can be carried out by attending to the visual modality alone (cf. Lemmens et al., 2001). Because of our interest in affective–computing research, we developed task–irrelevant emotionally–charged earcons that carried either a positive or a negative valence, using the major/minor distinction in Western tonal music (e.g. see Crowder, 1984, 1985a), to investigate whether affective–congruency effects are phenomena that researchers in affective computing should take into account. The present study, therefore, employs a stimulus–response compatibility paradigm to investigate whether maintaining affective correspondence is important for affective human–computer interfaces.

Earcons are audio messages used in human–computer interfaces to provide information and feedback. Earcons can include messages, functions, states, and labels (Blattner et al., 1989) and they are designed in a structured fashion starting by associating simple actions (e.g., "open") with simple earcons and by combining elementary actions into more complex actions (e.g., "open file" or "open folder") that are associated with compound earcons. Because of the hierarchical and structured nature of music, earcons usually are short musical fragments. Using typical musical transformation–dimensions like pitch, timbre, or rhythm (etc.; see Gaver, 1989; Pramana & Leung, 1999), elaborate families of earcons can be constructed that are structurally related to the functions they represent.

A transformation that seems missing in the literature, in our view, is to transform earcons into a variation in major or minor mode. Crowder (1984) discussed that musical pieces in major mode are always perceived as more positive whereas pieces in minor mode are perceived as carrying a more negative charge. Isolated chords, even when presented as briefly as 300 ms, also contain this stable positive/negative connotation. In the second study of a series of five, Crowder (1985a) backed these claims with empirical evidence. In our view, his evidence shows that the difference in affective appreciation of the major and minor mode can be incorporated in the set of transformations for earcons. The major/minor transformation can then be used specifically to create affectively–charged earcons for use in affective human–computer interfaces.

Using this distinction between major and minor mode, we were able to convert our multimodal picture–categorisation task from a task that originally contained only one affective aspect into a variation containing two affective features. We expected that the introduction of this additional affective component would lead to affective–congruency effects, because the newly added affective valence of the earcons could relate to the affective valence of the responses —participants had to respond using positively or negatively valenced Yes– or No–response buttons. The pattern that we expected was based on the dimensional overlap, which is the (degree of) similarity of a stimulus and response feature (Kornblum et al., 1990), between the affective valence of the earcons and that of the responses that participants had to execute. For instance, we expected that participants would be faster executing a positive response to trials that contained an earcon in major mode compared to trials (requiring a positive response) with an earcon in minor mode, because the former trial has a more natural relation between the affective aspects of stimulus and response compared to the latter trial. The relationship between major and minor earcons and the response categories Yes versus No presumed here presupposes a positive valence associated with Yes–responses and a negative valence associated with No–responses. Of course, this relationship may be questioned and said to be highly task–dependent. The response "No" to the question of whether one has cancer or not, for example, will certainly have a positive valence. In the present context and tasks examined, however, care was taken that Yes– and No–responses were linked to positive and negative valences, respectively. Therefore, we called the former type of trials congruent whereas the latter type was called incongruent. The expectations for trials requiring a negative response were similar. A negative response to a trial with an earcon in minor mode was congruent and therefore expected to show shorter response–time latencies than negative responses to a trial with an earcon in major mode (which was incongruent).

3.2 Method

We used an affective animate/inanimate decision of line–drawings of animals and musical instruments by instructing participants to press Yes–

and No–buttons in response to the question "Is the picture you see that of an animal? Yes/no". This way, pictures of animals were (implicitly) associated with "yes" responses.

We used a task manipulation to balance the design of the experiment to ensure that participants executed positive (yes) responses to the pictures of the animals in one part of the experiment as well as positive responses to the pictures of the musical instruments in another part. For instance, one instruction had an example which simply confirmed the abovementioned question of the instruction ("... For instance, if you see a rabbit, press the button with the label 'Yes' ... "). The other instruction reversed this response–to–stimulus assignment: "... For instance, if you see a rabbit, press the button with the label '*No*' ... ". Essentially, in this condition we forced participants to provide a wrong answer to the question.

3.2.1 Participants

Twelve psychology students took part in the experiment for either € 3,– or for partial fulfilment of course requirements. Their mean age was 23.1 years; three participants were male.

3.2.2 Materials, Stimuli & Design

Because of the multimodal aspect of the experiments both auditory and visual stimuli were employed. The auditory stimuli were the earcons in major and minor mode which were C4 and C5 major and minor triads in root position (see Crowder, 1985a) with a duration of 2500 ms, that were created by a professional sound–designer using a Roland midi module. The visual stimuli consisted of sixteen black and white line–drawings of animate and inanimate objects (see Fig. 3.1). The animate objects were limited to animals: a squirrel, a butterfly, a cat, a dog, a lion, a bird, a cow, and a frog. The inanimate objects (limited to musical instruments) comprised a violin, a trumpet, a drum, a saxophone, a guitar, a flute, an accordion, and a harp. Care was taken to ensure that all pictures were approximately of the same size when displayed on the experimental equipment.

The experiment was carried out on a Macintosh PowerMac G3 that was equipped with a 17 in. screen. A button box attached to the experimental computer was used to accurately synchronise the presentation of the visual and auditory stimuli, and to register response–time latencies.

Figure 3.1: Example pictures that we used in the present study. Note that the same pictures were also used in the other studies reported in this thesis that used the category of animals and musical instruments.

One button was labelled “yes” and another button was labelled “no” (the actual labels were in Dutch); the order of the labelling was counterbalanced between subjects to prevent effects of preferred hand and to prevent a confound due to a possible natural tendency to assign affirmative responses to the right hand (Wentura, 2000). Simple stereophonic–headphones were used to present the earcons (although no stereophonic effects were used).

The experiment employed a blocked within–subjects design with the factors Compatibility and Picture category. The design incorporated four conditions that were different with regard to the fixed relation between the affective valence of the earcons and the affective property of the response within a trial block (see Table 3.1). One trial block reflected an affect–congruent relation: all pictures of animals, requiring a “yes” response, with the sound of the C5 major chord, and all pictures of the musical instruments combined with the C5 minor chord; another trial block implemented the affect–incongruent relation (the pictures of animals with the minor chord and pictures of the musical instruments with the major chord). The remaining two blocks implemented a condition in which both categories of pictures were accompanied by the same sound (called affectively neutral). These neutral blocks were included as control conditions. Note that Table 3.1 also shows that we always called the combination of major earcons and positive responses (as well as minor earcons and negative responses) congruent whereas combinations of minor and positive (or major and negative) were always called incongruent.

The task manipulation ensured that the experimental design was balanced with respect to the assignment of responses (yes or no) to the two categories of (visual) stimuli (animals or musical instruments). Each of the four trial blocks was carried out once in each response–to–stimulus assignment. This resulted in 8 experimental blocks: Four trial blocks for the Animal⟶Yes response assignment and another four for the musical Instruments⟶Yes assignment. Half of the participants started with the Instrument⟶Yes task and then carried out the other assignment; the other half started with Animals⟶Yes.

Each trial block contained 32 stimuli: 16 stimuli in the above indicated picture–sound combinations and 16 without sound to create a baseline condition (see Table 3.1). The total number of trials therefore amounted to 32 trials × 4 blocks × 2 tasks = 256 trials. Within a block the stimuli were randomised differently for each stimulus list. The presentation order of the blocks over participants was according to a digram–balanced Latin

Table 3.1: Visualisation of the design of the experiment. The various combinations of pictures and types of sound are presented for both tasks.

Task: Animal ⟶ Yes

Picture category	Sound type			
Animal	**major (8)**	*minor (8)*	major (8)	minor (8)
	no sound (8)	*no sound (8)*	no sound (8)	no sound (8)
Instrument	**minor (8)**	*major (8)*	major (8)	minor (8)
	no sound (8)	*no sound (8)*	no sound (8)	no sound (8)

Task: Instrument ⟶ Yes

Picture category	Sound type			
Animal	*major (8)*	**minor (8)**	major (8)	minor (8)
	no sound (8)	**no sound (8)**	no sound (8)	no sound (8)
Instrument	*minor (8)*	**major (8)**	major (8)	minor (8)
	no sound (8)	**no sound (8)**	no sound (8)	no sound (8)

Note. Sound type represents the major/minor distinction used to create the earcons in major and minor mode, but also includes the no–sound baseline trials (n = 16) within a trial block (n = 32). Numbers in parentheses represent the number of trials. The trial blocks marked in bold comprised the congruent condition whereas italics indicates trial blocks that were called incongruent.

square.

3.2.3 Procedure

Participants were instructed to press buttons according to a question that was posed before commencing the experiment. For both tasks the question was "Is the picture you see that of an animal?". However, for the Instruments⟶Yes assignment, participants were instructed to press the "no" button to pictures of animals and the "yes" button upon presentation of a musical instrument using examples in the written and verbal instruction. Participants were instructed to do this quickly and accurately; they were not explicitly instructed to ignore the sounds, but neither were they encouraged to relate the sounds to the pictures they would see or to the responses they would make. Participants could practice on sixteen trials randomly drawn from the set of experimental trials.

The presentation of a single trial started with the presentation of a fixation cross for 500 ms (plus an alert beep). After a 500 ms pause, the visual and auditory stimuli were presented simultaneously. The visual stimulus was presented for 300 ms. The maximal response–time was set at 2500 ms. The inter–trial interval was set at 1000 ms. The entire experimental session had a duration of approximately 25 minutes.

3.3 Results

Before data analysis, response omissions and errors were pruned from the raw data. The number of trials removed amounted to 2% of the total number of trials of both tasks. Because of the small number of errors committed and because the number of errors did not differ between the two instructions, no error analysis was carried out. Because only the data from the congruent and incongruent blocks were of interest for the current study, the data from the affectively–neutral blocks were not used in the following statistical analysis; in a similar fashion the baseline trials (without sound) were also not included.

A repeated measures ANOVA was carried out with Picture category (animal or instrument) and Compatibility (affectively compatible or affectively incongruent) as within–subject factors. Means and standard errors are presented in Figure 3.2.

The analysis showed that participants were significantly faster when they responded to pictures of animals (444 ms) than when they responsed to musical instruments (504 ms; $F(1, 11) = 13.529$, $MSE = 3182.2, p = .004$). The main effect for Compatibility was significant ($F(1, 11) = 7.388$, $MSE = 2175.2$, $p = .020$). Participants responded, on average, 36 ms faster to incongruent trials than to congruent trials (456 ms and 492 ms, respectively).

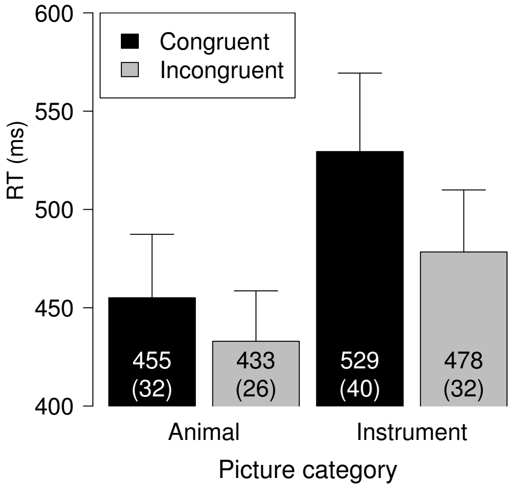

Figure 3.2: Mean response–time latencies (ms) of the overall analysis of Congruency and Picture category. RT's and SE's (in parentheses) at the bottom of each bar.

The interaction Picture category × Compatibility ($F(1, 11) = 8.322$, $MSE = 301.9$, $p = .015$) showed that the difference between the congruent and incongruent condition was significantly larger for the musical instruments (51 ms) compared to the same difference for the animal pictures (22 ms). That congruency effect proved to be significant for the musical instruments only ($F(1, 11) = 10.520$, $MSE = 2974.5$, $p = .008$).

We also carried out a post–hoc analysis to investigate how the somewhat awkward Instruments⟶Yes instruction affected performance. To do so, we decided to incorporate the different response–to–stimulus assignments as a factor Task (Animals⟶Yes or Instruments⟶Yes) in the analysis design Task × Picture category × Compatibility (see Table 3.2). A repeated measures ANOVA showed that participants were marginally faster in the Animals⟶Yes task (445 ms) than in the Instruments⟶Yes task (504 ms; $F(1, 11) = 4.517$, $MSE = 18579$, $p = .057$) but all interactions involving Task were not significant.

3.4 Discussion

Stimulus–response compatibility (SRC) phenomena have provided important insights for human–factors research, for instance concerning the spatial layout of interfaces (Vu & Proctor, 2003). Findings of semantic– and affective–compatibility effects (De Houwer, 1998; De Houwer & Eelen, 1998) seem to suggest that also on those levels of interface design, it is important to maintain correspondence to ensure optimal performance. Affective–compatibility effects are especially relevant for affective–computing research (Picard, 1997) but, to our knowledge, have not yet been investigated. We therefore carried out an experiment in which the auditory components (i.e., earcons) in a multimodal picture–categorisation task, that already employed affectively–charged positive and negative responses, were enriched with positively– or negatively–valenced affect. Based on the literature on affective–compatibility effects (e.g., De Houwer et al., 2001; De Houwer & Eelen, 1998), we expected that during information processing the relation between the affective charge of the auditory component of the stimulus and the affective valence of the responses would result in affective–congruency effects.

The results show a multimodal affective–congruency effect: Participants were significantly faster responding to certain combinations of affective valence of the earcons and affective valence of the responses. In

Table 3.2: Mean response–time latencies (ms) for the congruent and incongruent conditions for each picture category from each instruction with the data from the baseline condition included. Standard errors in parentheses.

<table>
<tr><th colspan="5">“Is the picture you see that of an animal?”</th></tr>
<tr><th colspan="5">Task: Animal⟶Yes</th></tr>
<tr><th></th><th colspan="4">Picture category</th></tr>
<tr><th></th><th colspan="2">Animal</th><th colspan="2">Instrument</th></tr>
<tr><th>Congruency</th><th>With earcon</th><th>Without earcon</th><th>With earcon</th><th>Without earcon</th></tr>
<tr><td>Congruent</td><td>433 (32)</td><td>432 (26)</td><td>495 (40)</td><td>458 (29)</td></tr>
<tr><td>Incongruent</td><td>412 (20)</td><td>413 (25)</td><td>440 (21)</td><td>428 (24)</td></tr>
<tr><th colspan="5">Task: Instrument⟶Yes</th></tr>
<tr><th></th><th colspan="2">Instrument</th><th colspan="2">Animal</th></tr>
<tr><th>Congruency</th><th>With earcon</th><th>Without earcon</th><th>With earcon</th><th>Without earcon</th></tr>
<tr><td>Congruent</td><td>564 (55)</td><td>453 (33)</td><td>478 (45)</td><td>502 (43)</td></tr>
<tr><td>Incongruent</td><td>517 (48)</td><td>474 (35)</td><td>455 (34)</td><td>442 (35)</td></tr>
</table>

Note. The data presented on one row represent the data acquired in one trial block. The column Without earcon presents the data of the baseline trials without sound. Note that the picture category assigned the Yes–response is always presented first. Congruency is determined by the dimensional overlap between the affective valence of the responses and the major or minor earcons. This implies that in the Animal⟶Yes task, in the congruent condition, the pictures of the animals were presented together with an earcon in major mode (also see Table 3.1). Also note that in the Instrument⟶Yes task, the data for the instrument pictures are presented first.

this experiment, the advantageous combinations were positive responses to trials with earcons in minor mode and negative responses to trials with earcons in major mode. Compared to our original expectations, however, this congruency effect shows a processing advantage for the less natural, incongruent trial types. If we assume that the positive response with an earcon in major mode (or a negative response with an earcon in minor mode) still is the more natural, congruent combination, then we must conclude that our participants interpreted the task and stimuli in a way that we did not anticipate.

We consider two potential explanations for the reversal of the congruency effect. Observe that the largest congruency effect was observed for the subset of data from the instrument pictures in the Instrument⟶Yes part. Considering the relative size of the effect in that subset of data compared to the effect in the other three subsets (see Table 3.2), the overall affective–congruency effect may have been overestimated due to the relatively large impact of the congruency effect of the instruments in the Instruments⟶YES part. Our first explanation relates to the presumed association between positive valence and Yes–responses and negative affective–valence and No–responses. Our request to execute the wrong response (relative to the implied answer to the question "Is it an animal?") in the Instrument⟶Yes part of the experiment may have uncoupled or even reversed the presumed association. That is, despite the fact that participants had to execute a No–response to the picture of an animal, this No–response may still have carried positive affect because of the implied positive response to the question posed in the instruction. This effect, secondly, may have been complicated further by the fact that participants may have relied on a stronger relationship between the pictures of the musical instruments and the earcons in major and minor mode, compared to that of the animals and the earcons, because the instrument pictures and earcons both have a musical character.

We also obtained an effect of Picture category representing a processing advantage of pictures of animals over the pictures of the musical instruments. In our view this processing advantage is similar to the animacy effect that von Studnitz and Green (2002) obtained in a language–switching task using the animate/inanimate distinction (see de Groot, 1990; Brousseau & Buchanan, 2004). The animacy effect refers to a simple processing advantage for biological living things, who's origin may lie in evolutionary benefits for animate objects (Caramazza & Shelton, 1998).

The findings carry two important messages for affective–computing research and human–factors research in general. The first message relates strongest to affective computing and is that affective–compatibility effects can cause performance decrements in affective human–computer interfaces that do not maintain affective correspondence between signals or events. The second message is that, for human–factors research, it is sometimes difficult to predict the strategy that participants formulate (often covertly) to carry out a task that they are given. The present results show that such effects of strategy can already be important in laboratory conditions; in real–life research conditions, their impact might even be stronger.

Chapter 4

Multimodal Affective–Congruency Effects: The Role of Judgemental Tendencies[†]

"...It is a riddle wrapped in a mystery
inside an enigma."

Sir Winston Churchill (1874–1965)
radio talk, October 1st, 1939

Abstract

This study aimed to investigate whether affective–compatibility effects that are reported in studies using visual linguistic stimuli could be replicated in a study of multimodal interference–effects involving audiovisual stimuli. To this end, we exploited in a picture–categorization task, in which participants were asked to discriminate between pictures of animals and inanimate objects, the affective connotations of the major and minor mode that exist in Western tonal music. The latter, task–irrelevant information accompanied the presentation of the pictures. A mixed between and within–group design was used in an attempt to capture the role of judgemental tendencies. The results show that the predicted affective–congruency effects were only present in the participant group answering a positive question ("Is this picture an animal?"). These effects were, however, reversed in the group answering a negative question ("Is this picture *not* an animal?"). Collectively, the results provide evidence for the involvement of judgemental tendencies as an important determinant of multimodal affective–congruency effects. We discuss the relevance of the present findings for the cognitive–ergonomic domain of affective computing.

[†] This chapter is based on the manuscript submitted as Lemmens, P. M. C., de Haan, A., and van Galen, G. P. (2004). *Multimodal Affective–Congruency Effects: The Role of Judgemental Tendencies.* Manuscript submitted for publication to Acta Psychologica.

4.1 Introduction

Research into congruency, correspondence, or, in general, compatibility effects has clearly demonstrated benefits for situations in which stimulus and response elements agree with a strong population stereotype (see Fitts & Deininger, 1954 and Fitts & Seeger, 1953). The majority of studies on stimulus–response compatibility (SRC) phenomena have been limited to relatively simple stimulus dimensions like location and color (Proctor & Reeve, 1990). Recently, however, researchers have taken up the challenge to investigate the effects of emotion and affect on task performance in an attempt to try and elucidate the representation of affect in human cognition. Because measuring affect directly is a rather complicated matter (involving measurements of the responses of the autonomous–nervous system, e.g., EEG, EMG, or the galvanic skin–response), or, when using questionnaires, is subject to personal interpretations of the participants, these researchers have turned to the indirect–measurement methods employed in SRC paradigms. They obtained, for example, affective Simon–effects (e.g., De Houwer et al., 2001; De Houwer & Eelen, 1998; Tipples, 2001; Voß et al., 2003) but affective Stroop–effects (Rothermund & Wentura, 1998) could not be found. There have also been several reports of affective–priming effects (e.g., De Houwer & Randell, 2004; Klauer & Musch, 2001; Moors & De Houwer, 2001; Wentura, 2000). Although strictly speaking the priming paradigm is not an SRC paradigm, the effects that were found typify the more general class of affective–compatibility effects (e.g. see Kornblum et al., 1990).

One aspect that seems to be missing in the literature on (general) affective–compatibility effects, to our knowledge, is information on multimodal affective–congruency effects (see Schriefers & Meyer, 1990, for a similar question regarding picture–word interference effects). Currently, most studies employ visually presented words that are both prime and target and each contain affective valence. In a multimodal variation of such a task, the task–relevant stimuli are presented, for instance, visually, whereas the task–irrelevant affective valence is presented auditorily. Our notion of multimodality, therefore, refers to a multimodal (auditory–visual) stimulus presentation and not to the cross–modal or intermodal dimensional overlap between the visually presented task–irrelevant stimulus feature and manually–executed response in, for instance, the affective Simon effect. Finding cross–modal interference effects in such a paradigm would not only corroborate existing evidence for affective–compatibility effects, but would also show that affectively–charged auditory material can interact with semantic information–processing. The latter is commonly assumed to mediate the processing of affective information (De Houwer & Randell, 2004; but see Wentura, 2000). The present study therefore used the affective connotation of the major/minor distinction in Western tonal music (Crowder, 1984; Gregory et al., 1996; Pittenger, 2003) to add an auditory component to a picture–categorization task in which participants had to discriminate between pictures of animals and inanimate objects. As in other affective–compatibility studies, the auditory affective–component (which we will refer to as auditory flankers in the present study) constitutes a task–irrelevant stimulus feature which will either correspond to or be in conflict with the affective property of the simple button–press responses that represented 'Yes'– and 'No'–answers to the question that was posed to the participants at the start of the experiment. In particular, in the present study, we exploited the affective connotation of the major and minor modes, which in Western music are known to be associated with positive and negative affect, respectively.

Our key prediction that a dimensional overlap (Kornblum et al., 1990) between the affective connotation of the major/minor chord distinction and the affective property of the responses that we asked our participants to generate, namely Yes– versus No–responses (see below), would systematically affect picture–categorization task performance, is based on two assumptions. The first assumption concerns the presumed affective connotation of the major/minor chord distinction. Crowder (1984) showed that the association of a positive mood, affective valence, and appreciation of musical works written in major mode and of a negative affective valence of music written in minor mode is a stable convention that musically–skilled as well as unskilled people recognize (Crowder, 1985b). In another study it was shown that children with an age of four years or older already attribute this conventional connotation to the major and minor modes (Kastner & Crowder, 1990), although infants of around 6 months of age do not yet seem to have acquired this association (Crowder et al., 1991). The distinction is present for chords that are presented in isolation and even when presented very briefly (Crowder, 1985a). Pittenger (2003) has recently confirmed (and extended) these findings. This evidence supports our assumption that chords in major mode are associated with positive affective–valence and that chords in minor mode are associated with negative valence and we expected that our participant group, that is, university students who usually have had (minimal) musical education and who regularly listen to Western tonal music, should be sufficiently sensitive to and appreciative of the emotional valence of isolated chords in major or minor mode (cf. Crowder, 1985a; Pittenger, 2003).

The second assumption on which we based our prediction that the dimensional overlap between the auditory flankers and the responses would affect performance of a picture–categorization task concerns the affective property of the Yes/No responses. We decided to use the Yes– or No–responses instead of the more common pronounciation of POSITIVE or NEGATIVE to add another response dimension to the aforementioned one and the previously employed response pairs of FLOWER/CANCER, NICE/NASTY, and COMEDY/CANCER (Tipples, 2001). To be able to determine what the more natural, congruent combinations of affective valence of the auditory flankers and that of the responses would be, we presupposed that Yes–responses are associated with positive valence and that No–responses are associated with negative valence. Of course, this relationship may be questioned and can be highly task–dependent. The response "No" to the question of whether one has cancer, for example, will certainly have a positive valence. Indeed, in a lexical–decision task a group of participants that Wentura (2000) instructed to respond with negative responses to word targets and positive responses to nonwords showed a reversed congruency effect (i.e., the mean response–time latency for the congruent condition was longer than that of the incongruent condition). In other words, Wentura (2000) found that his group of participants was slower to respond to a prime–target pair like SUMMER–HONEST despite the fact that this pair was affectively congruent. In his study, Wentura related his findings to the judgemental–tendency model proposed by Klauer and Stern (1992) in which the direction of affective–congruency effects is presumed to depend on the response assignment. We will return to this model–based interpretation of affective–congruency effects in the discussion section.

Based on these assumptions we examined in a picture–categorization task the performance of an affectively–congruent condition consisting of combinations of positive responses on the one hand and task–irrelevant chords in major mode on the other because they both have a positive valence. The same applied, of course, to combinations of minor chords and negative responses both having negative valence that were, therefore, included in the affectively–congruent condition as well. The affectively–incongruent condition comprised the opposite combinations: positive responses with minor chords and negative responses with major chords. We expected that participants would show shorter response times in the congruent condition of our picture–categorization task, in which they had to execute positive responses to pictures of animals and negative responses to pictures of inanimate objects by answering the question "Is the picture you see that of an animal?". Compared to the congruent condition, we expected that participants would show longer response–time latencies in the incongruent condition. We included a control condition, to complete the design, in which another group of participants was instructed to execute negative responses to the pictures of the animate objects. This group of participants was asked to answer the question "Is the picture *not* that of an animal?" to assign No–responses to the category of animate objects.

4.2 Method

4.2.1 Participants

Thirty–eight students (10 males) in psychology or cognitive science volunteered to participate in the experiment (mean age 23.5 years). They received € 4.50 or took part in the experiment for partial fulfillment of course requirements.

4.2.2 Materials, Stimuli & Design

Because of the multimodal aspect of the experiment both auditory and visual stimuli were employed. The auditory stimuli were C4 and C5 major and minor triads in root position with a duration of 2500 ms, that were created by a professional sound–designer using a Roland midi module. The visual stimuli consisted of 16 black and white line–drawings of animals and inanimate objects. The animate objects depicted were a squirrel, a butterfly, a cat, a dog, a lion, a bird, a cow, and a frog. The inanimate objects comprised a shoe, a bed, a balloon, a candle, a violin, a loaf of bread, a boat, and an airplane (see Fig. 4.1). These pictures were selected from two sets of pictures (of animals and inanimate objects) for which the mean affective rating, determined in a separate study with a different group of participants, did not differ statistically ($t(23) = 1.284$, *ns*). Care was taken to ensure that all pictures were of approximately the same size when displayed on the experimental equipment.

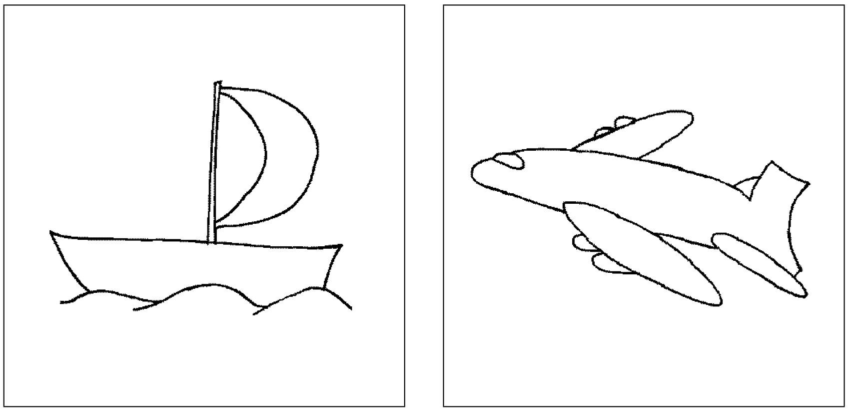

Figure 4.1: Two example pictures from the set of inanimate objects that we used in the present study. The pictures of animals were the same ones used in Chapter 3 (see Fig. 3.1).

The experiment was carried out on a Macintosh PowerMac G3 that was equipped with a 17 in. screen. A button box attached to the experimental computer was used to accurately synchronize the presentation of the visual and auditory stimuli and to register response–time latencies. One button was labeled "yes" and another button was labeled "no" (the actual labels were in Dutch); the order of the labeling was counterbalanced between subjects to prevent effects of preferred hand and to prevent a confound due to a possible natural tendency to assign affirmative responses to the right hand (Wentura, 2000). Simple stereophonic–headphones were used to present the auditory stimuli (although no stereophonic effects were used).

To balance response to stimulus–category assignments, we employed a between–subjects factor Task. Participants instructed to respond with Yes–responses to pictures of animals carried out the Animate–task and participants assigned to the group instructed to respond Yes to pictures of the inanimate objects carried out the Inanimate–task. For both groups each participant worked through the same within–subjects design (see Table 4.1) containing twelve trial blocks incorporating the factors Picture category (Animate or Inanimate) and Congruency (Affectively congruent or incongruent). The trial blocks were constructed using the following procedure that created two different sets of four trial blocks and another two sets of two trial blocks.

Individual trial blocks contained 32 stimuli: 8 pictures of animate

Table 4.1: Visualization of the design of the experiment. For each group of participants the various combinations of pictures and types of sound are presented.

Task: Animate ⟶Yes				
Picture category	Sound type			
Animate	**major (8)**	*minor (8)*	major (8)	minor (8)
	no sound (8)	*no sound (8)*	no sound (8)	no sound (8)
Inanimate	**minor (8)**	*major (8)*	major (8)	minor (8)
	no sound (8)	*no sound (8)*	no sound (8)	no sound (8)
Repetitions	4	4	2	2
Task: Inanimate ⟶Yes				
Picture category	Sound type			
Animate	*major (8)*	**minor (8)**	major (8)	minor (8)
	no sound (8)	**no sound (8)**	no sound (8)	no sound (8)
Inanimate	*minor (8)*	**major (8)**	major (8)	minor (8)
	no sound (8)	**no sound (8)**	no sound (8)	no sound (8)
Repetitions	4	4	2	2

Note. Sound type represents the major and minor auditory flankers, but also includes the no–sound baseline trials (n = 16) within a trial block (n = 32). Numbers in parentheses represent the number of trials. The two leftmost trial blocks were each replicated four times whereas the two rightmost trial blocks were each replicated twice. The trial blocks marked in bold are called congruent trial blocks whereas italics indicates trial blocks that are called incongruent.

objects and 8 of inanimate objects in the indicated picture–sound combinations (see Table 4.1) and the same 16 pictures without sound to create a baseline condition and to prevent participants from exploiting the fixed relation between picture and sound present in the congruent and incongruent trial blocks. Given a set of twelve different trial blocks, the total number of trials amounted to 32 trials × 12 blocks = 384 trials per experiment. Within a trial block the stimuli were randomized differently for each stimulus list. The presentation order of the trial blocks over participants was according to a Latin square.

4.2.3 Procedure

Participants were instructed to press one of two buttons to answer a question that was posed at the start of the experiment. Participants who were (randomly) assigned to the Animate–task were instructed to answer the question “Is the picture you see that of an animal?” by pressing a button labeled “yes” or “no” which implicitly associated positive responses to pictures of animals. Participants that were (randomly) assigned to the Inanimate–task were instructed to answer the question “Is the picture you see *not* that of an animal?”. This question implicitly associated negative responses to pictures of animals. Participants were instructed to do this quickly and accurately; they were not explicitly instructed to ignore the sounds, but neither were they encouraged to relate the sounds to the pictures they would see and the responses they would make. Participants could practice on ten trials randomly drawn from the set of experimental trials.

The presentation of a single trial started with the presentation of a fixation cross for 500 ms (plus an alert beep). After a 1000 ms pause, the visual and auditory stimuli were presented simultaneously. The visual stimulus was presented for 300 ms. The maximal response–time was set at 2500 ms. The inter–trial interval was set at 1500 ms. The entire experimental session had a duration of approximately 35 minutes.

4.3 Results

Before the statistical analyses the data were pruned from errors and omissions. The repeated–measures ANOVA on Task (between–subjects factor; Animate–task or Inanimate–task), Picture category (Animate or Inanimate), and Congruency (Affectively congruent or Affectively incongruent) was, therefore, carried out on the correct trials only. Besides an ANOVA by subjects (computing averages for 38 subjects over the 16 different pictures), we carried out an item analysis (computing averages for 16 items over 38 subjects). In this analysis, Picture category was a between–subjects variable. The ANOVA by subjects is reported as F_1 (or t_1) and the item analyses are reported as F_2 (or t_2).

4.3.1 Errors

The errors comprised only 1.2% of all available data points. A repeated measures ANOVA on error percentages from the data of the congruent and incongruent trial blocks only showed that participants committed more errors in the Inanimate– than the Animate–task (i.e., 8.2% and 1.3%, respectively; $F(1, 36) = 44.672$, $p < 0.001$). This effect, however, did not interact with effects of Picture category or Congruency. The latter factors did not show main effects on error percentages.

4.3.2 Reaction time

The mean response–time latencies (and associated standard errors) are presented in Figure 4.2. Contrary to our expectation, we did not find a main effect of Congruency. Pooled across the two participant groups, the mean RT's were 427 ms and 428 ms for the affectively–congruent and incongruent conditions, respectively ($F_1 < 1$; $F_2(1,28) = 1.796$, *ns*).

However, we did obtain an interaction between Congruency and Task ($F_1(1,36) = 11.615$, $p < 0.01$; $F_2(1,28) = 14.144$, $p < 0.001$). This interaction showed that participants in the Animate–task responded, as expected, faster to the congruent condition (430 ms) than to the incongruent condition (444 ms; $F_1(1,36) = 7.254$, $p < 0.05$; $t_2(15) = -3.394$, $p < 0.01$). Participants in the Inanimate–task showed the opposite pattern. The −13 ms difference between the congruent and incongruent condition (425 ms and 412 ms, respectively) was significant in the subject analysis and proved to be a weak trend in the item analysis ($F_1(1,36) = 4.671$, $p < 0.05$; $t_2(15) = 1.873$, $p = 0.081$).

We found that participants responded significantly faster to the pictures of animals than to the pictures of inanimate objects (410 ms and 445 ms, respectively; $F_1(1,36) = 75.562$, $p < 0.001$; $F_2(1,28) = 63.240$, $p < .001$). For Picture category we did not find an interaction with Task or with Congruency (F_1's < 1; F_2's < 1). The three–way interaction Task × Picture category × Congruency was not significant ($F_1 < 1$; $F_2(1,28) = 1.207$, *ns*). The results showed that the two task variations were not different with respect to response–time latencies ($F_1(1,36) = 1.436$, *ns*; $F_2(1,28) = 15.381$, $p < 0.001$).

4.4 Discussion

This study was carried out to investigate whether affective–compatibility effects that are reported in the literature (e.g., De Houwer et al., 2001, 2002) can be replicated in a multimodal version of a picture–categorization task. To do so we used the connotation that exists in Western tonal music that associates the major mode with positive affective–valence and the minor mode with negative valence (Crowder, 1984; Pittenger, 2003) and created

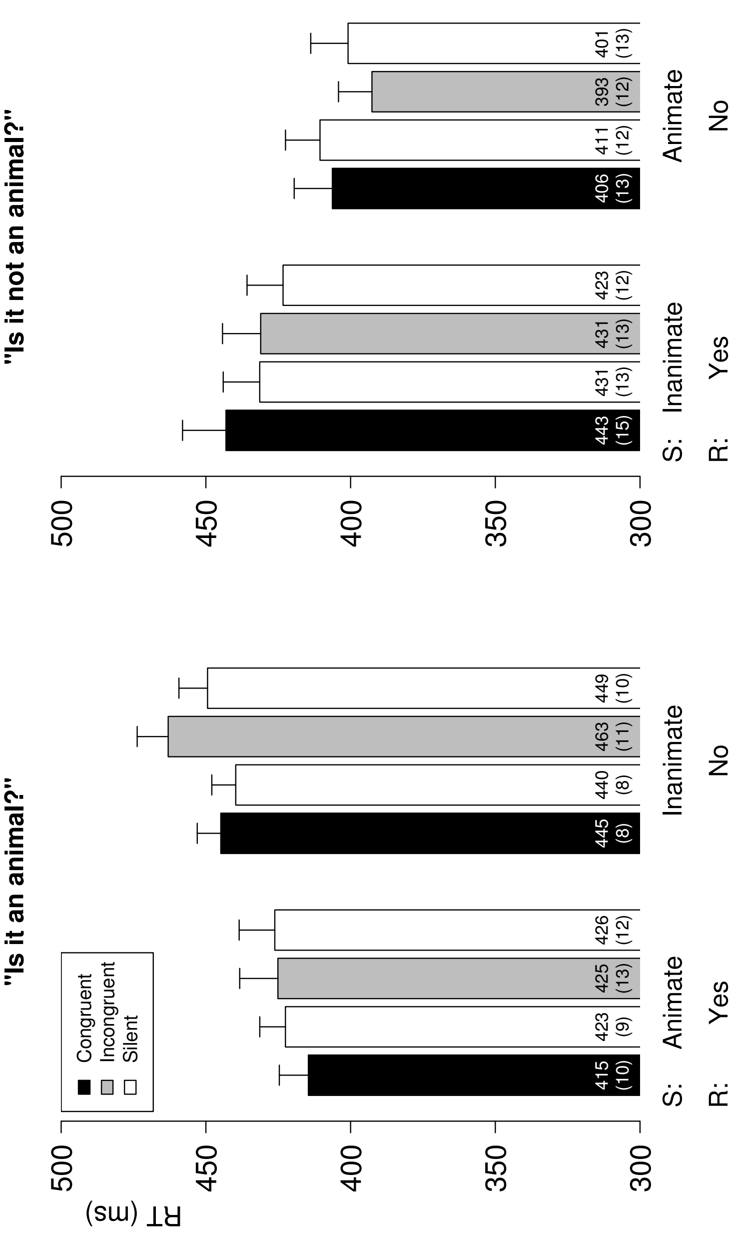

Figure 4.2: Mean response–time latencies (ms) for the congruent and incongruent conditions for animate and inanimate pictures with the baseline (silent) condition included. RT's and SE's (in parentheses) at the bottom of each bar. The pairs of black or gray and white bars reflect overall means of data points from the same condition. For instance, the black and white bars in the left panel for the animate objects reflect the means of the congruent condition and the associated baseline trials from the same condition (see Table 4.1). The directly neighboring gray and white pair reflects the data points from the incongruent condition. The picture categories are ordered according to the Yes– and No–response assignments: In each panel the picture category on the left was assigned the Yes–response. S indicates stimulus category; R indicates the instructed response.

a dimensional overlap between this affective connotation[1] and the affective properties of the responses that participants had to execute. Finding differences in response–time latencies to the congruent and incongruent conditions would be an interesting addition to the existing literature which, to date, has only provided evidence for affective–compatibility effects in the visual modality. A multimodal affective–congruency effect involving the auditory modality would also demonstrate that affective information presented in this modality is processed in the semantic system which has been proposed to mediate the processing of affective information (De Houwer & Randell, 2004).

The results of the present experiment were quite ambiguous. To our surprise, we did not find an overall difference between the affectively congruent and incongruent conditions. However, we did obtain an interaction of Congruency and Task that demonstrated that the participants carrying out the Animate–task showed the predicted congruency effect (i.e., shorter RT's to the congruent condition and longer RT's to the incongruent condition) whereas the participants in the Inanimate–task showed a reversed congruency effect with longer RT's to the congruent condition and shorter RT's to the incongruent condition. It seemed as if some aspect of the Inanimate–task completely reversed the pattern of overlap between the affective valence of the auditory flankers and the affective property of the responses. Initially we reconsidered the affective charge of the visual targets. Although the mean affective charge of the (super)sets of the visual targets did not differ significantly, the ratings for some pictures employed in the present study were relatively high. The affective charge of these pictures, however, did not modulate our main finding of the Task × Congruency interaction as is explained in Appendix 4.B.

Comparing a spreading–activation account and a judgemental–tendency model (Klauer & Stern, 1992) of affective priming, Wentura (2000) found a similar second–order interaction between the manipulation of the response assignments and the congruency manipulation in his study of the affective–priming effect. He took this interaction as evidence for the involvement of judgemental tendencies in the affective–priming effect, because of the spreading–activation account and the judgemental–

---

[1] Note that, despite converging evidence, obtained by Crowder (1985a, 1985b), Gregory et al. (1996), and Pittenger (2003), on the positive and negative affective charge of the major and minor mode, respectively, we carried out a rating study (see Appendix 4.A) investigating the affective charge of the auditory flankers that we employed. In brief, this study confirmed our assumptions.

tendency model only the latter predicts that response assignments are important in determining which conditions are affectively–congruent and which conditions are affectively–incongruent.

The judgemental–tendency theory by Klauer and Stern (1992), that was proposed to understand how attitudes guide memory–based judgements, revolves around a 3–step process[2] proposing that, first, the affective features of stimuli are automatically extracted and evaluated. If there happen to be two affective features, both are extracted and their congruency is evaluated. In the second step, the evaluation results in an affirmative response–tendency if the stimulus features are affectively congruent and a tendency to reject the relation if it is incongruent. For instance, for the prime–target pair SUMMER–HONEST the affective charges are congruent, and the response–tendency would be to confirm this relationship. For the pair SUMMER–THIEF, on the other hand, the response tendency would be to reject the relation. In a controlled process, the third step, this tendency is then verified as being appropriate or inappropriate in relation to the instructed response.

It is this step that generates the predictions concerning the interaction of congruency and response assignment. If, for instance, participants are instructed to execute a negative response to a certain target, say, negative responses to nouns, this negative response is congruent (and thus shorter RT's are expected) with a prevailing response–tendency to reject the affectively *in*congruent prime–target pair SUMMER–THIEF. Note that this procedure of comparing a response tendency and the instructed response is similar to the logical–recoding model for SRC effects that was proposed by Hedge and Marsh (1975). In this model, the same recoding rule ('identity' or 'reversal') is applied to both stimulus features (S_r and S_i; Lu & Proctor, 1995). Faster responses are expected when the rules for the relevant and irrelevant stimulus dimensions are equivalent (i.e., either identity or reversal) compared to a stimulus for which the rules are different (identity vs. reversal, or v.v.). Hedge and Marsh (1975) favored a horse–race model to account for the differences in response–time latencies. This model was formalized by de Jong et al. (1994) and extended by Ridderinkhof (2002).

Although our study and that by Wentura (2000) employ different tasks

---

[2] Klauer and Stern (1992) did not incorporate the automatic extraction and evaluation of the affective information as a separate step in the logic of their model, as the title of their paper shows. Wentura (2000) mentions the stage of extracting the affective information as the additional step.

and a different (temporal) order of presentation of prime/flanker and target, the studies also have some similarities. Both experiments employed an emphasis on a particular category of stimuli and both studies manipulated the response assignment to that category. For these reasons and the comparable pattern of results, we argue that judgemental tendencies may have been involved in the realization of the multimodal affective–congruency effect involving audiovisual stimuli in the present experiment. A side–effect of the instruction to the participants in the Inanimate–task may have contributed to the reversal of the congruency effect that we found, namely that participants use the animate category as reference category because that leads to the simplest response–selection strategy. For instance, in the Inanimate–task we initially considered the picture of a boat with a major chord as affectively congruent, because the instructed response to the question "Is it not an animal?" is positive, which is congruent with the affective valence of the chord in major mode.

However, the instructed way to determine the correct answer to the question posed is quite complex, so participants may have inserted an additional step in their response–selection strategy to circumvent the complex negation. They may have selected the appropriate response following the procedure "NO, this boat is not an animal, so I have to press the Yes–button" instead of "YES this boat is not an animal". The additional processes to handle the complex logic inherent in the question posed in the Inanimate–task, however, do not correspond to the affirmative response–tendency based on the affective relation between the auditory flanker and the instructed response. Thus longer latencies were observed for the condition that was a priori labeled as congruent (see Fig. 4.2). If the chord had been in minor mode, the response tendency would have been to reject the affective relation between auditory flanker and instructed response which would have been congruent with the internal comparison to the reference category ("NO, not an animal") resulting in shorter latencies for the incongruent condition. Such a response–selection strategy is not necessary in the Animate–task, however, because its instruction was much simpler and our original expectations regarding multimodal affective–congruency effects can therefore be maintained.

So instead of a simple affective–congruency relation that we expected between auditory flanker and instructed response, we found that the multimodal affective–congruency effect actually is the result of a judgemental tendency that is related to an additional set of mental processes that participants probably used to realize a simpler response–selection strategy in the Inanimate–task in which the instructed process to determine the response was quite complex. The absence of a main effect of Task, reflecting the between–subjects response–assignments manipulation, corroborates this conclusion.

Besides the congruency and task effects, we also obtained a consistent processing advantage for the animals. Such processing advantages have frequently been found in other studies that employ the animate versus inanimate distinction (e.g., Brousseau & Buchanan, 2004; de Groot, 1990; von Studnitz & Green, 2002). Findings in patient studies have shown that specific brain lesions can differentially affect the capability to discriminate between animate and inanimate categories. Caramazza and Shelton (1998) reasoned that this specific cortical development might be related to the importance of the animate category during the evolution of human cognition. Another idea is that the animacy effect is inherently closely related to affective processing (Brousseau & Buchanan, 2004). An alternative explanation of the Picture–category effect may be the difference in class homogeneity of the animate and inanimate pictures, the former category consisting of a much smaller and more homogeneous set than the latter category (see Lemmens, de Haan, van Galen, & Meulenbroek, 2004a). However, the presently found evidence for the role of judgemental tendencies in multimodal affective–congruency does not rely on this Picture–category effect.

To summarize, our attempt to elicit a multimodal affective–congruency effect using the affective connotation of the major/minor distinction was successful as indicated by the significant differences between the affectively congruent and incongruent conditions in both tasks. This finding is important because it shows that affective–information processing does not depend on the visual modality (which one might conclude based on the existing literature) and that affective–compatibility effects can come about in a cross–modal fashion. The reversal of the effect in the Inanimate task hints at the involvement of judgemental tendencies in the realization of the congruency effect. The multimodal affective–congruency effect also shows that affectively–charged auditory information is processed by the semantic system. So despite the fact that our auditory flankers were task–irrelevant and did not have a clear semantic connotation, they were nevertheless processed by the semantic system and their affective features still “meddled” (Wentura, 2000) with the formation of the responses.

Finally, we would like to mention that the present findings are important for affective computing (Picard, 1997) which is a human–factors research initiative to investigate the merits of including aspects of human emotion to improve the efficiency of human–computer interfaces. The finding that the affective connotation of the major/minor distinction can be used to transmit affect using the auditory modality can be readily used by sound–designers that need to create affectively–charged earcons (see Blattner et al., 1989). The affective–congruency effects show that it is important to maintain affective correspondence to ensure that performance stays at an optimal level. Moreover, the reversal of the multimodal affective–congruency effect in the Inanimate–task demonstrates that complex instructions can lead to unexpected choices by participants that can interfere with the predetermined patterns of congruency.

4.A A rating study: Affective charge of the auditory flankers

Although the literature on the perception and appreciation of the major and minor consistently attributes a positive appreciation (happy) and a positive affective charge to the major mode and a negative appreciation (sad or dreamy) and negative affective charge to the minor mode (see Crowder, 1984, 1985a, 1985b; Gregory et al., 1996; Pittenger, 2003), we nevertheless carried out a rating study to investigate the affective charge of the auditory flankers that we employed in the present study.

A group of naive participants was instructed to mark each sound they heard as negatively or positively charged on a 5–point scale on which position 3 was explicitly labeled as representing a neutral position (a sound with no affective charge present) and a score of 5 represented a positive affective charge. In addition to the auditory flankers that we used in the present study, we also presented flankers that we employed in other studies and each sound was presented a number of times to obtain repeated measures of the affective rating.

We expected a significant difference in rating between flankers in major mode and flankers in minor mode with the major mode receiving a higher score on the scale than the flankers in minor mode (Crowder, 1985a, 1985b). We also expected (overall) higher ratings for chords from a higher pitch class than for chords from a lower pitch class (see Collier & Hubbard, 2001, 2004).

4.A.1 Method

Participants

Forty–four participants took part in the experiment for partial fulfillment of course requirements. Their mean age was 21.5 years and 14 participants were male.

Materials

Ten different sounds that were different on three aspects: Mode (major or minor mode), Duration (300 ms, 1250 ms, or 2500 ms), and Height (high, C5 octave, or low, C4 octave). Four out of the ten sounds were the sounds that we used in the main study of the present chapter: the major or minor chords of 2500 ms in a high or low variation. A similar set of four sounds was used that had a duration of 300 ms. The sounds with a duration of 1250 ms were only available in the high variation. Each sound was presented four times and the presentation order was pseudo randomized.

Procedure

The participants were seated in similar rooms compared to the participants in the experiment on the multimodal affective–congruency effect. They were instructed to determine, by their personal preference, whether the sound(s) that they would hear carried a positive or negative affective charge. The instruction explicitly stated that participants should use the third point of the 5–point scale for sounds that they felt were neither positively charged nor negatively charged.

There were no practice trials and participants had around 10 seconds to determine the rating before a fixation cross and alert beep indicated the imminent presentation of the next sound (that would start 1 second after the alert beep). The participants wrote down the ratings on unmarked sheets of paper with 40 lines numbered lines, one for each sound, with the numbers 1–5 preprinted equally spaced over the line.

4.A.2 Results & Conclusions

Visual inspection of the raw data did not highlight clear deviations from the normality assumption for an ANOVA. We therefore computed a mean affective rating for each flanker for each participant and carried out a repeated–measures ANOVA on Mode × Height × Duration. Note that we did not include the flankers with a duration of 1250 ms in the analysis because doing so would have resulted in an incomplete design as these flankers were only available in the high (C5) variation. The mean ratings are presented in Table 4.2.

A repeated–measures ANOVA confirmed that the rating of the flankers in major mode was significantly higher ($F(1,43) = 38.776$, $MSE = .429$, $p = .000$) than the rating of the flankers in minor mode (3.21 vs. 2.77). As expected, we found that higher flankers were, in general, rated more positive (3.60) than flankers that were an octave lower in pitch (2.43; $F(1,43) = 36.992$, $MSE = 2.99$, $p = .000$).

The main effect of Duration was not significant ($F < 1$) but Duration did modulate Mode (Mode × Duration: $F(1,43) = 5.262$, $MSE = .147$,

Table 4.2: Mean affective rating (on a 5–point scale with a score of 5 representing a positive affective charge) for the major and minor chords that were employed in the present study as well as in other studies. Standard errors are presented in parentheses.

		Mode	
Height		Minor	Major
High (C5)			
	300 ms	3.2 (0.08)	3.7 (0.09)
	2500 ms	3.4 (0.17)	3.9 (0.17)
Low (C4)			
	300 ms	2.4 (0.09)	2.7 (0.10)
	2500 ms	2.0 (0.17)	2.6 (0.17)

p = .027): the difference in affective rating is larger for the 2500 ms flankers (0.53; $F(1, 43) = 39.230$, $MSE = .521$, $p = .000$) than for the shorter 300 ms flankers (difference, 0.34; $F(1, 43) = 27.583$, $MSE = 1.782$, $p = .000$). So, although the difference in rating of the 300 ms flankers is significant (a result previously obtained by Crowder, 1985a, 1985b), the increased duration of the 2500 ms flankers results in a clearer perception as positively or negatively charged for the flankers in major and minor mode, respectively.

The interaction of Duration and Height just reached the level of significance: $F(1, 43) = 4.319$, $MSE = .806$, $p = .044$. The 2500 ms sounds were rated more positive in the C5/high variation than the 300 ms flankers (3.65 vs. 3.45, respectively) whereas the situation was reversed for the lower C4 flankers (2500 ms, 2.33; 300 ms, 2.53). The 3–way interaction as well as the interaction between Mode and Height were not significant ($F < 1$ or $p > .1$).

Posthoc comparisons of the ratings for each pair of flankers in major and minor mode (e.g., the 300 ms C4 flankers) showed that all flankers in major mode were consistently rated higher on the 5–point scale than the accompanying flanker in minor mode and that all differences between these scores were significant (all p's < .001, Bonferroni corrected).

From these results we conclude that our assumption, that the sounds that we employed were affectively charged and, more specifically, that the flankers in major mode were rated more positive than the flankers in minor mode, was not unwarranted. Therefore, we conclude that the presumed and confirmed affective charge of the auditory flankers can be used to, a priori, determine the affective compatibility of the relationship between the flankers and other affectively–charged task–features. From this line of reasoning we gather that the affective charge of the flankers was not involved in the reversal of the affective–congruency effect that we observed for the group of participants that carried out the Inanimate–task in the present study. That is, it is unlikely that the different task–requirements for the Inanimate–task resulted in a different (affective) appreciation of the auditory flankers.

4.B A rating study: Affective charge of the visual targets

The multimodal affective–congruency effect that we report in the present study builds on the dimensional overlap (Kornblum et al., 1990) of the affective charge of the auditory flankers and the affective charge of the responses with the added assumption that the visual targets should be affectively neutral because otherwise an overlap between the task–relevant stimulus and the response or between the task–relevant (S_r) and task–irrelevant stimulus (S_i) features can potentially confound with the affective–congruency effect that is under investigation. It is, however, nearly impossible to find such a set of visual targets for our participant population as nearly every picture (i.e., nearly every object or concept) has the potential to evoke some kind of affective response and this response can vary strongly between participants.

We therefore investigated whether the mean affective charge of the category of animals and inanimate objects was significantly different. To do so we created a rating questionnaire in which all available visual targets were presented and we requested the participants to rate each picture on a 7–point scale by its perceived affective charge. We expected that the mean affective rating of these categories would not be different, although both categories could be rated as either positive or negative (as compared to the affectively neutral midpoint of 4).

4.B.1 Method

Participants

On a voluntary basis, we selected 25 participants (8 males) that had not recently taken part in any of our studies. With the exception of their gender, the participants could fill in the questionnaire anonymously.

Materials

The questionnaire consisted of handout in which 30 pictures were presented (6 on each page). The pictures comprised 12 pictures of animals (including the 8 pictures that we used in the present study), 8 pictures of musical instruments and 10 pictures of inanimate objects (again including the 8 pictures employed in the present study). Space for writing

Table 4.3: Mean affective rating (and SD's in parentheses) on a 7–point scale (with 7 representing a positive valence) for the picture categories employed in the present study as well as the other studies reported in this dissertation.

	Picture category		
	Animals	Inanimate objects	Musical instruments
Rating	4.67 (0.44)	4.56 (0.36)	4.44 (0.50)

down the names as well as a 7–point scale for each individual picture was provided.

Procedure

Participants were handed the questionnaires and were requested to fill the requested data at a leisurely moment at which they could not be disturbed. For each picture the participants had to fill in the name of the picture and their personal affective rating on a 7–point scale of which position 4 was implicitly labeled as affectively neutral.

4.B.2 Results & Conclusions

Before statistical analysis, the data from one participant were removed because of procedural inconsistencies. We verified visually whether we could see clear deviations from the normal distribution. Because this was not the case, we computed mean affective ratings for each category of pictures for all participants and subjected these scores to paired t–tests to investigate whether the mean ratings were statistically different. We limit the discussion to the pair of categories of most interest: animals vs. inanimate objects. The mean affective ratings for all categories are presented in Table 4.3.

The paired t–tests showed that the 0.11 difference between the category of animals and inanimate objects was not significant ($t(23) = 1.284$, *ns*). However, closer inspection showed that the subset of pictures that we employed in the present study contained a number of pictures of animals that were rated relatively high whereas the inanimate objects contained pictures that were rated relatively low. Comparing the mean affective rating for these 8 pictures of animals (4.86) and the 8 pictures of inanimate objects (4.47) showed that these ratings were different: $t(23) = 3.899$, $p = .001$.

Therefore, we removed the highest rated pictures from the set of animals (i.e., the squirrel and the butterfly; see Figure 4.3) and the pictures rated lowest from the set of inanimate objects (i.e., the loaf of bread and the shoe) and again tested the difference between the (new) mean rating for animals (4.63) and the rating for the inanimate objects (4.60). This 0.03 difference proved non–significant ($t(23) = 0.300$, *ns*).

We then carried out the same repeated–measures ANOVA on Task × Picture category × Congruency, that we used in the main analysis, on a subset of data from which the pictures of the squirrel, butterfly, the shoe, and the loaf of bread were removed. The effects of interest are the (absence of a) main effect for Congruency (Affectively congruent, 427 ms; Affectively incongruent 427 ms; $F < 1$) and the interaction with Task ($F(1, 36) = 11.608$, $MSE = 714.6$, $p = .002$). In the Animal–task we again observed the predicted congruency effect (congruent, 428 ms; incongruent, 443 ms; $F(1, 36) = 6.565$, $MSE = 2858.4$, $p = .015$) whereas the effect reversed in the Inanimate–task (congruent, 425 ms; incongruent, 411 ms; $F(1, 36) = 5.188$, $MSE = 2858.4$, $p = .029$). These results are equivalent to the findings in the present study in which the data from all visual targets were used.

If one assumes that a score of around 5.5 for the picture of the squirrel and the butterfly represents a genuine positive valence and one also assumes that participants completely ignore the auditory flanker, then alternative compatibility relations between the affective charge of these specific pictures and the affective property of the responses can be determined. For instance, for the picture of the squirrel (with a positive affective valence) this would imply that all trials acquired from the participants carrying out the Animal–task can be considered to be affectively congruent, because throughout that task a Yes–response (also with a positive affective charge) was associated to pictures of animals. On the other hand, in the Inanimate–task No–responses (with a negative affective charge) were associated to pictures of animals. Thus, in the data from this task, trials involving the picture of the squirrel can be considered affectively incongruent because the (momentarily assumed) positive affective charge of the squirrel is incongruent with the affective valence of the instructed response. We therefore computed mean RT's for the Animate–task (congruent condition) as well as Inanimate–task (the incongruent condition) on all data points without sound from all blocks,

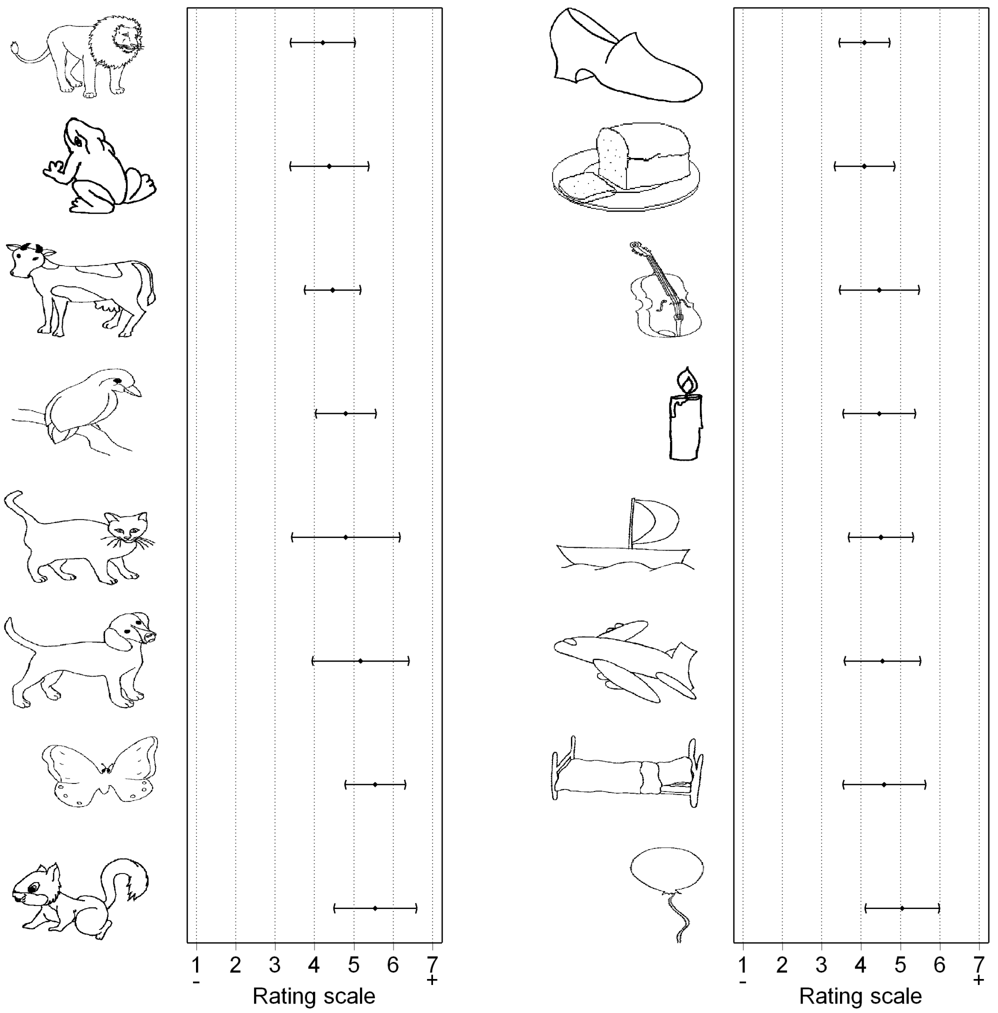

Figure 4.3: Mean affective rating for each picture employed in the present study, per category in an increasing order of rating. The whiskers represent the standard deviation. On the abscissa is the rating scale with ratings above 4 indicating a positive valences and ratings below 4 representing negative valences. Mean rating for the pictures of animals was 4.67 and the mean rating for the inanimate objects was 4.56. Note that some of the pictures are slightly deformed due to rescaling for this figure.

for each subject, for the pictures of the squirrel and the butterfly. We then assessed whether the difference between these sets of data was significant using an independent–samples t–test. This test showed that the difference was not significant ($t(36) = 1.400, p > .17$). We conclude that, if present, the affective charge of the pictures of the squirrel and the butterfly did not result in compatibility relations with the affective charge of the responses. Note, however, that side–effects of the blocked design may invalidate these conclusions and that removing the assumption of the ignored flankers would complicate this line of reasoning considerably.

We conclude, tentatively, that the affective charge of the pictures that were rated relatively high (low) compared to the other pictures of animals (inanimate objects) did not influence the realization of the multimodal affective congruency effect. In general, this applies to the entire set of pictures that we employed in the present study.

Chapter 5

Effects of Animacy and Categorical Homogeneity in Multimodal Affective Categorization†

> Dealing with failure is easy: Work hard to improve. Success is also easy to handle: You've solved the wrong problem. Work hard to improve.
>
> ---
>
> *Alan Perlis' Epigrams*

Abstract

In the present study we evaluate the extent to which the relative homogeneity of picture categories elicits dedicated information–processing strategies that facilitate reaction speed when people perceive and identify stimuli that belong to a specific category. In a study of the multimodal affective–congruency effect Lemmens, de Haan and Van Galen (2004) concluded that the observed response–time advantage for one of the picture categories that they used could be attributed to a specific processing advantage favoring animates over inanimates. The results from the present study, however, show that categorical homogeneity plays a more important role in the picture–category effect, as well as in demonstrating the multimodal affective–congruency effect itself, than an effect of animateness. These findings prompt for a re–interpretation of the results reported by Lemmens et al. (2004).

[†] This chapter is submitted as Lemmens, P. M. C., de Haan, A., van Galen, G. P., & Meulenbroek, R. G. J. (2004). *Effects of Animacy and Categorical Homogeneity in Multimodal Affective Categorization.* Manuscript submitted for publication to Psychological Research.

5.1 Introduction

The animacy effect refers to a processing advantage for animals compared to inanimate objects. For instance, Lawrence (1971, cited in Pashler, 1998a) found that people more readily detected an animal name in a list of non–animal names than a non–animal name in a list of animal names. With lexical decision and animateness–categorization tasks, de Groot (1990) observed that associative–priming effects were larger for animateness categorization than for lexical decision, but only when the targets were animate (Exp. 1–2). Brousseau and Buchanan (2004) obtained an animate advantage that they attributed to a possible association between those animate items and subjective valence. In an animate/inanimate categorization within a language–switching task, von Studnitz and Green (2002) observed faster animate decisions than inanimate decisions. Such overall processing advantages were also observed by de Groot (1990). It is known that the animate/inanimate distinction serves a special role in category development in children (e.g. see, Mandler, 1997; Rakison & Poulin–Dubois, 2004). Patient studies have shown that lesions can differentially affect animate and inanimate categories (Caramazza & Shelton, 1998). These findings all highlight a special role for animate objects in the formation and recognition of categories of objects.

Recently, Lemmens et al. (2004) reported a multimodal affective–congruency effect in an animate/inanimate picture–categorization task. In that task, we presented major and minor chords simultaneously, as auditory flankers, with the visual targets. Crowder (1985a) had shown that a stable connotation existed for major chords related to positive valence and minor chords to negative valence and we exploited this connotation to create a dimensional overlap between the affective feature of the responses (Yes– or No–button presses) and the auditory flankers. We found, however, that the overlap did not exist between the auditory flankers and the overtly executed Yes– and No–responses but between the flankers and an internal covert response (e.g., "YES, this is an animal picture, so, by instruction, I now have to press the No–button") that was realized by emphasizing the importance of the animate category in the instruction. The dimensional overlap between the covert responses and the auditory flankers resulted in a set of congruent flanker–response pairings (e.g., YES and major chord or NO and minor chord) and pairings that were incongruent (e.g., YES and minor chord or NO and major chord). The response–time differences between the congruent and incongruent conditions were found to be significant, indicating a multimodal affective–congruency effect. Besides this effect, we also found a significant effect of picture category showing a processing advantage for the pictures of animals. Because we obtained this advantage when participants had to respond with Yes to the animates as well as when the animates were assigned a No–response, we concluded that an animacy effect was the most likely cause since a speed advantage for Yes–responses alone could not account for the findings.

However, the emphasis that was put on the animals may in itself have prompted strategic use of that emphasis, possibly also resulting in processing advantages for animals, because participants may have (subconsciously) allocated more resources to the processing of the emphasized category. Processing advantages due to the increased resource allocation could thus have been confounded with a processing advantage due to the animacy effect. Moreover, the set of animate objects was a homogeneous well–defined set almost exclusively consisting of mammals. Literature on categorization (e.g. see Rosch & Lloyd, 1978) has shown that items in such sets are often categorized by exemplar or prototype similarity which is a fast and efficient process. Categorization by exemplar similarity, however, was impossible for the set of inanimate objects because this set was rather heterogeneous including pictures of means of transportation, footwear, and a toy. The inanimate objects could, therefore, not be tagged with a unique simple tag like "toy" or "tool". We propose that a more laborious and slower hypothesis–testing approach to categorization was required to verify each item against a relatively large superset of inanimate objects. As the choice of categorization type is the result of the differences in degree of homogeneity, the difference may have caused a processing advantage for the animate objects because of their increased homogeneity. This processing advantage is again confounded with the animacy effect.

Therefore, in our view three accounts exist that can all explain the effect of picture category that Lemmens et al. found: (a) an effect of animacy, (b) processing differences due to the differences in homogeneity of the categories of animate and inanimate objects, and (c) strategic use of the accentuation of the animate objects. The present study was set up to disentangle these different accounts. The implicit encouragement of strategically using the emphasis on the animate objects to realize the internal covert response stage was part of the experimental design of our previous experiments that we did not want to change to maintain a degree of comparability between the results. We thereby attempted to extricate the animacy effect and potential processing differences due to differences in categorical homogeneity.

Hence, we sought to decrease the heterogeneity of the category of inanimate objects. In contrast to the earlier experiments we exploited the two homogeneous categories of animals and musical instruments as animate/inanimate categories, rather than a homogeneous set of animals in combination with a heterogeneous set of miscellaneous inanimate objects. This setup enabled us to test the contrast between a potential processing advantage for the pictures of animals that was entirely due to their animate nature versus a processing advantage that was inadvertently created as result of the different degrees of homogeneity. If we were correct in surmising the involvement of the animacy effect, changing the inanimate objects to a homogeneous set of musical instruments should not matter for the effect of picture category. That is, a response–time advantage for pictures of animals should still be found regardless whether the second category would be a well defined homogeneous category like musical instruments (or toys or tools) or an ill–defined heterogeneous category. If, on the other hand, processing strategies due to differences in homogeneity were responsible for the effect of picture category, decreasing the differences between the degree of homogeneity of each category of objects should result in similar categorization strategies and thus in similar response–time latencies for both categories.

Note, however, that we maintained the same emphasis on one specific category of objects that Lemmens et al. used. As we explained above, in itself this accentuation may have had the effect of increased allocation of resources to the processing of the emphasized category, resulting in decreased response times. If this would be the case in the present study, response–time differences should be observed in a group of participants for which one category is emphasized and the other one is not. For instance, if one group participants would receive emphasis on the set of musical instruments, for that group a processing advantage would be expected for the category of instruments compared to the category of animals (v.v. for emphasis on animals for another group of participants). Thus, within a group of participants an effect of picture category would still be expected, favoring the category of pictures receiving accentuation in the instruction.

Also observe that if categorical homogeneity indeed proves to be responsible for the picture–category effect, it may affect the multimodal affective–congruency effect itself because earlier results already showed larger congruency effects for the heterogeneous inanimate objects. The change in categorical homogeneity of the inanimate objects might therefore have an impact on the multimodal affective–congruency effect in the present study. With two homogeneous categories of objects, the congruency effect was expected to shrink in absolute size.

5.2 Method

5.2.1 Participants

Forty–nine participants (13 male, mean age 22.1 years) took part in the experiment for either € 4.50 or partial fulfillment of course requirements.

5.2.2 Materials, Stimuli & Design

Because of the multimodal aspect of the experiments, both auditory and visual stimuli were employed. The auditory stimuli were C4 and C5 (i.e., a lower and higher variant of) major and minor triads in root position with a duration of 2500 ms, that were created by a professional sound designer using a Roland midi module. The visual stimuli consisted of sixteen black and white line-drawings of animals and musical instruments (see Fig. 3.1). The animals depicted were a squirrel, a butterfly, a cat, a dog, a lion, a bird, a cow, and a frog. The musical instruments comprised a violin, a trumpet, a drum, a saxophone, a guitar, a flute, an accordion, and a harp. Care was taken to ensure that all pictures were approximately of the same size when displayed on the experimental equipment.

The experiment was carried out on a Macintosh PowerMac G3 that was equipped with a 17 in. screen. A button box attached to the experimental computer was used to accurately synchronize the presentation of the visual and auditory stimuli, and to register response–time latencies. One button was labeled "Yes" and another button was labeled "No" (the actual labels were in Dutch); the order of the labeling was counterbalanced between subjects to prevent effects of preferred hand and to prevent a confound due to a possible natural tendency to assign affirmative responses to the right hand (Wentura, 2000). Simple stereophonic headphones were used to present the auditory stimuli (although no stereophonic effects were used).

To counterbalance the emphasis on each category of pictures, we employed a between–subjects factor Task. Participants instructed to respond with Yes–responses to pictures of animals carried out the Animal task, whereas participants assigned to the group instructed to respond Yes to pictures of musical instruments carried out the Instrument task. For both groups each participant worked through the same within–subjects design (see Table 5.1) containing twelve trial blocks incorporating the factors Picture category (Animals or Instruments) and Mode (Major or Minor). The blocks were constructed using the following procedure that created two different sets of four trial blocks and another two sets of two trial blocks.

The twelve trial blocks were different with regard to the fixed relation between the Mode of the auditory flanker and the Picture category it was associated to. For both tasks, in four trial blocks all pictures of animals were combined with the sound of either the C4 or C5 major chord, and all pictures of musical instruments were combined with the minor chords. Similarly, in another four trial blocks the pictures of animals were combined with either the C4 or C5 minor chord and pictures of the instruments with either the C4 or C5 major chord. The remaining four trial blocks implemented a condition in which both categories of pictures were accompanied by the same sound (two trial blocks with only major chords, C4 or C5, and two with only minor chords).

Individual trial blocks contained 32 stimuli: 8 pictures of animals and 8 of musical instruments in the indicated picture–sound combinations (see Table 5.1) and the same 16 pictures without sound to create a baseline condition and to prevent participants exploiting the fixed relation between picture and sound especially present in blocks 1 and 2. Given a set of twelve different trial blocks, the total number of trials amounted to 384 trials per experiment. Within a trial block the stimuli were randomized differently for each stimulus list. The presentation order of the trial blocks over participants was according to a latin square.

5.2.3 Procedure

Participants were instructed to press buttons according to a question that was posed before commencing the experiment. Participants assigned to the Animal categorization task were instructed to answer the question "Is the picture you see that of an animal?" by pressing a button labeled "Yes" or "No'. For the group of participants that carried out the Instrument categorization task, the question was "Is the picture you see that of a

Table 5.1: Visualization of the design of the experiment. For each group of participants the various combinations of pictures and types of sound are presented.

Task: Animal ⟶ Yes				
Picture category	Sound type			
Animal	**major (8)**	*minor (8)*	major (8)	minor (8)
	no sound (8)	*no sound (8)*	no sound (8)	no sound (8)
Instrument	**minor (8)**	*major (8)*	major (8)	minor (8)
	no sound (8)	*no sound (8)*	no sound (8)	no sound (8)
Repetitions	4	4	2	2
Task: Instrument ⟶ Yes				
Picture category	Sound type			
Animal	*major (8)*	**minor (8)**	major (8)	minor (8)
	no sound (8)	**no sound (8)**	no sound (8)	no sound (8)
Instrument	*minor (8)*	**major (8)**	major (8)	minor (8)
	no sound (8)	**no sound (8)**	no sound (8)	no sound (8)
Repetitions	4	4	2	2

Note. Sound type represents the major/minor distinction of the factor Mode, but also includes the no–sound baseline trials (n = 16) within a trial block (n = 32). Numbers in parentheses represent the number of trials. The two leftmost trial blocks (in the text referred to as blocks 1 and 2) were each replicated four times whereas the two rightmost trial blocks (referred to as blocks 3 and 4) were each replicated twice. The trial blocks marked in bold can be considered congruent blocks whereas italics indicates trial blocks that can be considered to be incongruent.

musical instrument?”. Participants were instructed to do this quickly and accurately; they were not explicitly instructed to ignore the sounds, but neither were they encouraged to relate the sounds to the pictures they would see and the responses they would make. Participants could practice on ten trials randomly drawn from the set of experimental trials.

The presentation of a single trial started with the presentation of a fixation cross for 500 ms (plus an alert beep). After a 1000 ms pause, the visual and auditory stimuli were presented simultaneously. The visual stimulus was presented for 300 ms. The maximal response time was set at 2500 ms. The inter–trial interval was set at 1500 ms. The entire session had a duration of approximately 35 minutes.

5.2.4 Data analysis

We analyzed the data using the factor Mode to further investigate possible differences between trials with auditory flankers in major mode and trials with flankers in minor mode. Because Mode was nested within congruency we were unable to include congruency as a factor in the analysis design and for each task we therefore assessed the difference between animal–major and animal–minor pairings (similarly for the instrument pairs) using paired samples t–tests as an alternative way to evaluate congruency.

In a similar way, the balancing of the assignment of responses to each picture category that was carried out using the factor Task —participants carrying out the Animal task always responded Yes to animal pictures and No to instruments, whereas this relation was reversed for the participants assigned to the Instrument task— created a nested design, preventing the inclusion of the factor Response (Yes or No) into the ANOVA. Therefore, provided there was no effect of Task, we used independent–samples t–tests to assess differences in Yes– and No–responses within a picture category.

Besides an ANOVA on subjects (computing averages for 49 subjects over the 16 different pictures), we carried out an item analysis as well (computing averages for 16 items over 49 subjects). In this analysis, Picture category was a between–subjects variable. The item analyses are reported as F_2 (or t_2), whereas the ANOVA on subjects is reported as F_1 (or t_1). All statistical tests were carried out against a significance level of $\alpha = .05$.

5.3 Results

We earlier only used data from blocks 1 and 2 for statistical analysis because the congruency effect was of main interest. Here we will also report on the data from blocks 3 and 4 that can be seen as variants of blocks 1 and 2 without the fixed relation between picture and sound that is present in the latter blocks. We decided to incorporate the baseline trials within each block in the figures that we have drawn but not to take these data points into the statistical analyses because a comparison of trials with and without sound seemed inappropriate.

Because we employed two homogeneous categories of pictures, we did not expect a difference in mean response–time latencies between the category of animals and musical instruments. As the present design enabled us to disentangle the effects of degree of homogeneity versus processing advantages due to speeded Yes–responses, we expected that there would be no response–time differences for Yes– and No–responses to each picture category. On the contrary, we expected an interaction of Task and Picture category that would speak for strategic use of the emphasis on one of the categories of objects (animals in the Animal task and musical instruments in the Instrument task). Finally, the earlier experiments showed the strongest congruency effects in the heterogeneous category of inanimate objects. Because we changed that category, we anticipated less pronounced congruency effects in this experiment.

Errors

Before analysis, errors and omissions were removed from the set. These comprised only 2.1% of the total number of data points and were therefore not analyzed further.

Response times

The Task × Picture category × Mode repeated–measures ANOVA (see Fig. 5.1) on blocks 1 and 2 showed no main effect for Picture category. Mean response–time latencies for animals were 439 ms and 442 ms for the musical instruments ($F_1 < 1$; $F_2(1,28) = 2.043$, $MSE = 92.1$, *ns*). The between–subjects factor Task did not show a main effect as well (Animal task, 438 ms; Instrument task, 443 ms; $F_1 < 1$).

The interaction of Task and Picture category, that we anticipated, was significant (see Fig. 5.1, $F_1(1,47) = 28.605$, $MSE = 1377.1$; $F_2(1,28) = 141.658$, $MSE = 92.1$). The interaction revealed that for the Animal task mean RT's for pictures of animals were faster (422 ms) than those for pictures of instruments (454 ms, $F_1(1,47) = 18.562$, $MSE = 5508.3$; $t_2(14) = -10.542$) whereas this was reversed for the Instrument task (Animals, 456 ms; Instruments, 431 ms; $F_1(1,47) = 10.670$, $MSE = 5508.3$; $t_2(14) = 6.759$).

However, this result was not due to the assignments of Yes–responses to the emphasized category. Comparisons of the mean RT's for Yes–responses and No–responses for each category of objects across participant groups showed that although participants in the Animal task (animal pictures with Yes–responses) were 34 ms quicker compared to participants executing No–responses to animal pictures (associated to

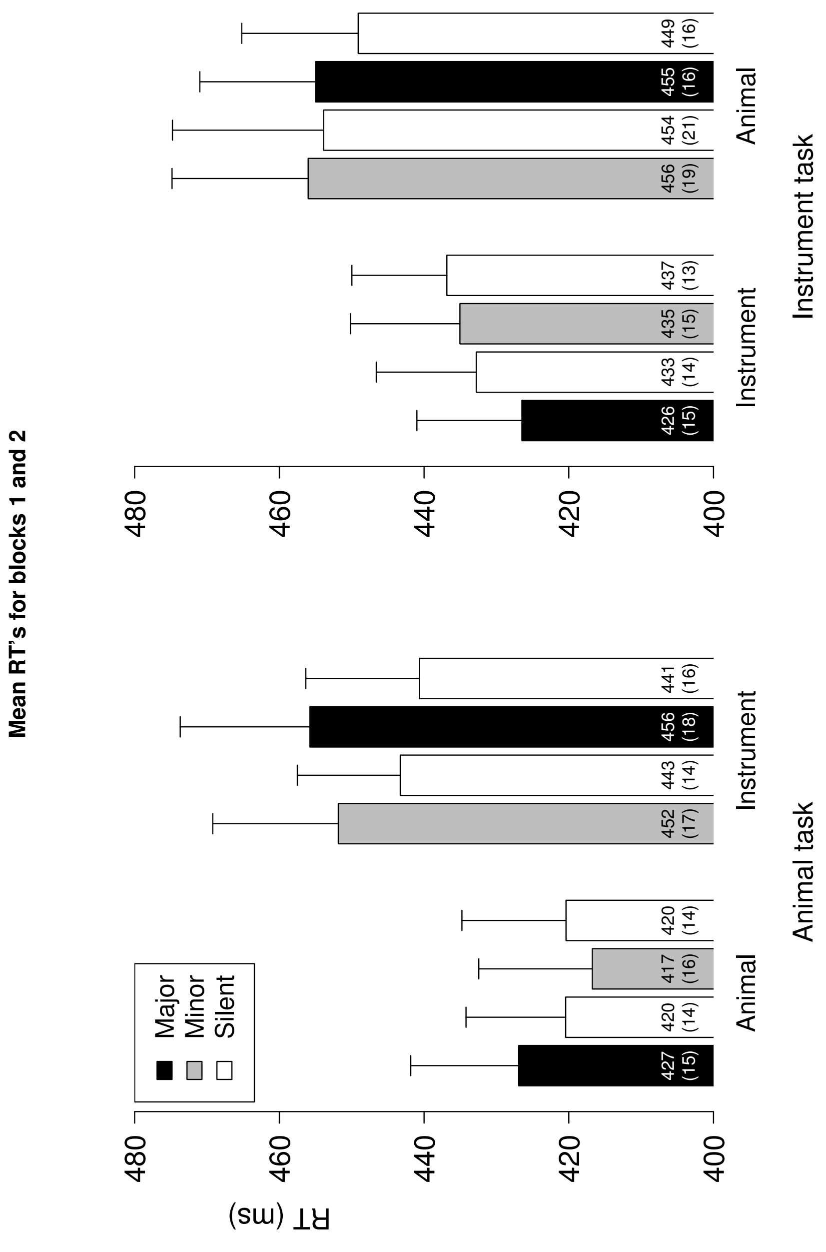

Figure 5.1: Mean response–time latencies (ms) from blocks 1 and 2 for each task for both picture categories and major and minor mode, including baseline trials from the same block. RT's and SE's (in parentheses) at the bottom of each bar. The pairs of black or gray and white bars reflect means of data points from the same block for each category of pictures. For instance, the black and white bars in the left panel for the animal pictures reflect the means for the animal–major combinations from blocks 1 and the associated baseline trials from the same block (see Table 5.1). The directly neighboring gray and white pair reflects the data points from blocks 2. The picture categories are ordered according to the Yes– and No–response assignments: In each panel the picture category on the left was assigned the Yes–response.

the Instrument task), this difference was not significant ($F_1(1,47) = 2.142$, $MSE = 6481.8$, *ns*). Similarly, for Yes– and No–responses to pictures of instruments, the 23 ms advantage for Yes–responses was not significant ($F_1(1,47) = 1.052$, $MSE = 6185.5$, *ns*).

The data from blocks 3 and 4 showed a similar pattern of results (see Fig. 5.2) with non–significant main effects for Picture category (Animal pictures as well as Instruments, overall, 444 ms; $F_1 < 1$) and Task ($F_1 < 1$; Animal task, 442 ms; Instrument task, 445 ms).

For the blocks 3 and 4, the interaction Task × Picture category was significant ($F_1(1,47) = 49.890$, $MSE = 929.8$). In the Animal task, participants were quicker to respond to pictures of animals (427 ms) than to pictures of instruments (458 ms; $F_1(1,47) = 25.860$, $MSE = 3719.2$) whereas participants were quicker responding to pictures of instruments (430 ms v.s. 461 ms; $F_1(1,47) = 24.069$, $MSE = 3719.2$) in the Instrument task. Again these response–time advantages were not due to the response assignments. Yes–responses to pictures of animals (427 ms) were not statistically different from No–responses (461 ms; $F_1(1,47) = 1.825$, $MSE = 7761.0$, *ns*). Similarly, Yes–responses to Instrument pictures (430 ms) did not deviate from the No–responses (458 ms; $F_1(1,47) = 1.294$, $MSE = 7174.9$, *ns*).

In the data from blocks 1 & 2, the main effect of Mode was significant ($F_1(1,47) = 5.693$, $MSE = 61.6$, $F_2(1,28) = 5.571$, $MSE = 103.2$) indicating that participants were quicker in responding to minor chords (438 ms) than to major chords (443 ms). Mode did not interact with Task ($F_1 < 1$, $F_2 < 1$) nor with Picture category ($F_1 < 1$, $F_2 < 1$). The three–way interaction Task × Picture category × Mode was not significant in the main analysis ($F_1 < 1$) nor in the item analysis ($F_2(1,28) = 2.134$, $MSE = 103.2$, *ns*).

Figures 5.1 and 5.2 show the effect of Mode, favoring minor chords, in all combinations of picture category and sound in both tasks. Note that only for the picture category that received emphasis in the instruction (associated with Yes–responses), the difference between major and minor (reflecting an effect of congruency) approaches levels similar to those obtained in the earlier study. However, in that study the largest congruency effects were actually observed for the category that was *not* accentuated. Analyses on the data from blocks 1 & 2 showed that although in the Animal task mean response times to animal pictures with major (427 ms) were slower than the same pictures with a minor chord (417 ms), this difference was not significant ($t_1(24) = 1.639$, $SEM = 6.20$, *ns*); the item analysis, however, showed a significant difference ($t_2(7) = 4.746$,

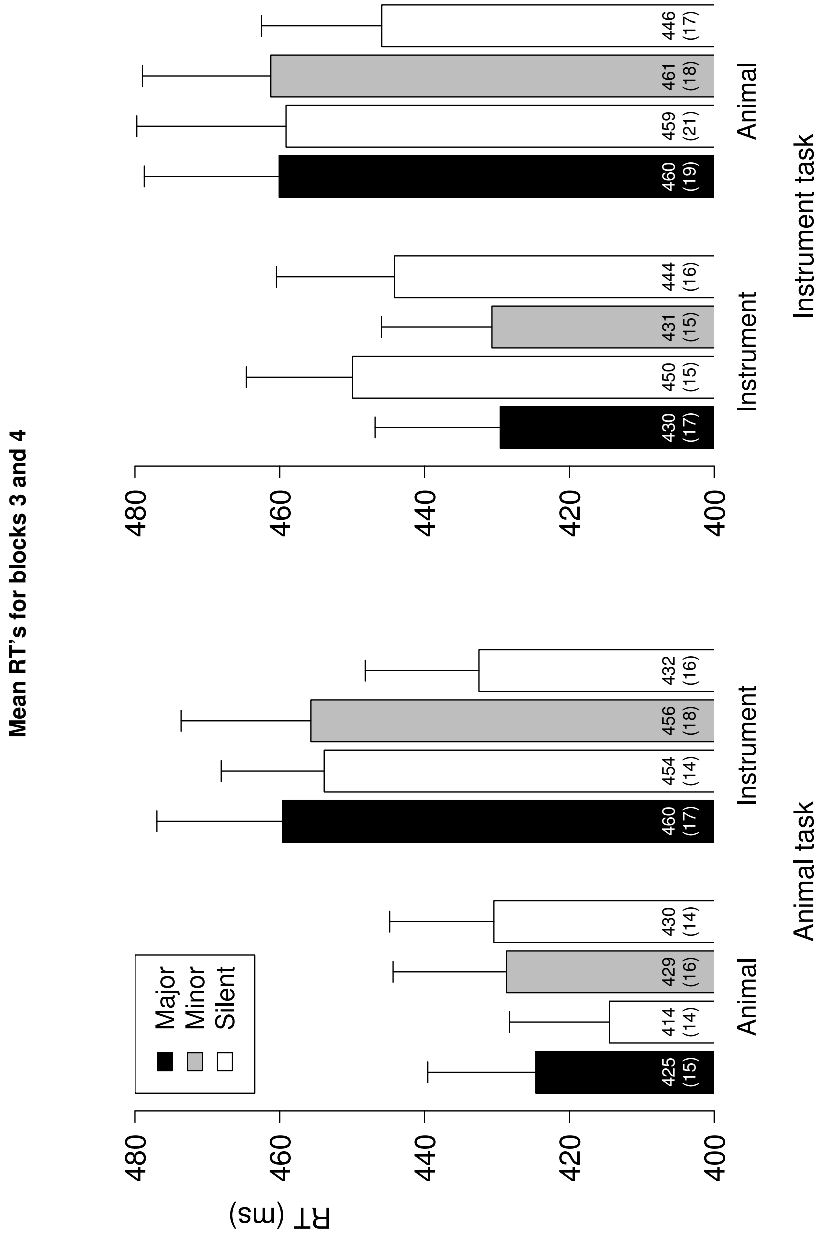

Figure 5.2: Mean response–time latencies (ms) for each task, for both picture categories and major and minor mode using the data from blocks 3 and 4. RT's and SE's (in parentheses) at the bottom of each bar. The pairs of silent trials and trials with sound refer to data points from one block. For instance, for the animal pictures in the left panel, the mean for major sounds and baseline trials are from blocks 3 and the directly neighboring gray and white pair of means is from blocks 4 (see Table 5.1). The picture categories are ordered according to the Yes– and No–response assignments: In each panel the picture category on the left was assigned the Yes–response.

SEM = 2.22). The difference, in the Instrument task, between major (435 ms) and minor pairings (426 ms) with the instruments was almost the same as the 10 ms difference observed between animal–major and animal–minor pairings in the Animal task, but as before, this difference was not significant ($t_1(23) = 1.548$, $t_2(7) = 1.783$, both *ns*).

5.4 Discussion

The present study revealed interesting findings, particularly regarding the picture–category effect that was obtained in our earlier study. There we found consistent response–time advantages for the category of animals and we reasoned that this advantage for the animals was mostly likely due to an animacy effect, specifically, because we still found the advantage for the animate objects when participants had to execute No–responses to these objects which excluded the possibility that the effect of picture category was due to response–time differences between Yes– and No–responses. Post hoc analyses of our earlier study showed that the picture–category effect was significant for the congruent condition ($F(1, 48) = 38.004$, $MSE = 858.5$, $p = .000$) as well as the incongruent condition ($F(1, 48) = 29.019$, $MSE = 1030.4$, $p = .000$). We therefore had clear indications that an animacy effect, also observed in several other lines of research employing the animate/inanimate distinction (e.g., Brousseau & Buchanan, 2004; de Groot, 1990; von Studnitz & Green, 2002), was a likely cause of the picture category effect.

However, this explanation referring to an effect of animacy was confounded with two other possible explanations. First, it may have been that participants strategically employed the emphasis that was put on the category of animate objects. Although the emphasis was an explicit experimental manipulation, it may have caused the processing advantage for the animates nonetheless because it might have triggered allocating additional resources to processing that category. Secondly, a difference in degree of homogeneity existed between the set of animate and inanimate objects that were employed in our earlier study. Nearly all animate objects that we used were mammals whereas the inanimate objects were much more diverse, consisting of means of transportation, footwear and other objects. Participants may have used fast and efficient exemplar or prototype categorization to classify the animate objects and may have required the slower hypothesis–based testing (Rosch & Lloyd, 1978) to categorize the inanimate objects. The differences between these categorization processes may also have caused a processing advantage for animate objects. Both explanations are confounded with the explanation using the animacy effect and cannot be teased apart using the data of Lemmens et al. (2004).

Because the encouragement to focus on one specific category was an essential experimental manipulation, in the present study we alleviated the differences in degree of homogeneity by using a well–defined set of inanimate objects, namely musical instruments, instead of the heterogeneous set employed in the earlier study. We predicted that this change would result in the removal of the effect of picture category because both sets of pictures could now be categorized using exemplar or prototype based classification. The results of the present study clearly show no effect of picture category, thus confirming our prediction. This conclusion was corroborated by a similar finding for the blocks of trials in which the fixed relation between picture and sound was removed (blocks 3 and 4). We also predicted that the picture category effect was not due to a difference in processing speeds for Yes– and No–responses. The data, from blocks 1 and 2 as well as 3 and 4, again confirm this prediction.

Finally, we predicted that processing advantages due to the emphasis on one specific category of objects would result in an interaction of Task and picture category, reflecting processing advantages for animals in the Animal task and for instruments in the Instrument task. The data clearly confirm this prediction. Thus, the present study falsifies our earlier conclusion that the picture category effect was due to an effect of animacy. Instead, the data suggest that the emphasis on a specific category of objects resulted in the response–time advantage. Moreover, in the original data the processing advantage for the animals was confounded with a response–time advantage due to different categorization processes for animate and inanimate objects.

Thus, the data gathered in the present study propose that degree of homogeneity of two categories employed in an experiment can create processing advantages reflected in response–time differences for one category that are indeed due to the distance in homogeneity and not due to a specific characteristic of one of the categories. This conclusion may have consequences for other studies employing the animate/inanimate distinction in which processing advantages for animate objects frequently have been found.

For instance, in a naming study Brousseau and Buchanan (2004) used pictures of animals, fruits and vegetables, weapons and instruments, and tools and clothes. They found that “animals were associated with the fastest mean vocal RT” (p. 245), and that, in general, naming times for the biological category (animals, fruits, and vegetables) were faster than those for the non–biological category. They proposed a category distinction with a biological advantage (although only in young, healthy females). However, inspection of their stimulus material, in our view, shows that the non–biological category is less homogeneous than the biological category. Although a naming task involves different mental processes than the choice response–time tasks employed in the present and earlier studies, that difference in degree of homogeneity may nevertheless cause similar processing strategies in the naming task employed by Brousseau and Buchanan (2004) as those found in the present study.

Such detailed information of the stimuli that were used is not available for the studies by de Groot (1990) and von Studnitz and Green (2002). However, inspection of example material that de Groot (1990) presents in Table 5.1 (p. 108) seems to show rather heterogeneous categories for both the animate as well as the inanimate objects. The comparable heterogeneity will most likely not have caused differential processing strategies for animate and inanimate objects. In a similar, although opposite way, von Studnitz and Green (2002) seem to have prevented strategies by using relatively well–defined homogeneous categories like fruits and toys (see Table 2, p. 245).

Regarding the multimodal affective–congruency effect that was the focus of our original investigation, the present data suggest that also for this effect the degree of homogeneity plays an important role. We already observed larger differences between the congruent and incongruent condition for the inanimate objects. For instance, the regular affective categorization task in our original study showed a difference of around 10 ms between the congruent and incongruent condition for the animate objects (quite similar although in opposite direction to the 10 ms difference between major and minor chords for animal pictures in the present Animal task), whereas this difference doubled to approximately 20 ms for the inanimate objects (which is clearly not the case in the present Animal task). Their reversed task showed a similar pattern although smaller in number. In the present study, with two homogeneous sets of objects, no congruency effects (expressed as the difference between major and minor chords for each level of Task × Picture category) were found however.

The only relevant difference between the present and earlier study is the degree of homogeneity of the inanimate objects and, therefore, the processes used to categorize the different sets of objects (i.e., exemplar–based vs. hypothesis–based categorization). We speculate that the hypothesis–testing approach to categorization takes more information into account (to improve the quality of a decision) than the exemplar–based categorization used for the animate objects. The fixed relation between picture and sound within blocks 1 & 2 could well have been that additional information. Because this fixed relation reflects the effect of congruency, we observed larger congruency effects for the set of objects categorized by hypothesis–based testing (i.e., the inanimate objects). The relatively low number of errors that the participants committed corroborates the claim of high quality responses. We tentatively conclude that the multimodal affective–congruency effect observed in earlier studies (e.g. Lemmens et al., 2004, a study that was based on research reported in Van Esch–Bussemakers, 2001) may have been overestimated. The increased response–time differences between major and minor chords in the hypothesis–based categorization of the inanimate objects were possibly caused by the strategic use of the fixed relation (within one block) between picture and sound. Awaiting further evidence, the present findings lead us to, at least, question the explanation of Lemmens et al. (2004) that was purely in terms of an affective overlap between the covert responses (realized by emphasizing a particular category) and the valence of the major and minor chords.

An additional difference with the earlier study is that in the present study we observed an effect of Mode: participants responded faster to minor chords than they did to major chords. Although we then could not explicitly report on response time differences between major and minor chords (resp., 437 ms and 435 ms) because of the nesting in congruency, a post hoc analysis with Mode instead of Congruency showed that this difference was not significant (Exp. 2, Task × Picture category × Mode. Main effect of Mode, $F(1, 48) < 1$). This difference in the presence and absence of an effect of Mode can be explained referring to the nesting of Picture category and Mode in the factor Congruency. Compare, for instance, the Animal task from the present study and the regular affective–categorization task from our earlier study. In the latter study the congruency effect for the animals essentially reflects response–time advantages for major chords, because animals required Yes–responses that were affectively congruent with the major chord and the congruent condition was faster than the incongruent condition. For the inanimate objects, however, the situation was reversed. Inanimate objects requiring a No–response were affectively congruent with a minor chord, therefore associating minor chords with faster responses and major chords with slower responses. Collapsed over picture category, response–time differences between major and minor chords therefore cancelled out. As can be seen in Figure 5.1, in the present study however, the advantage for the minor chords in the Animal pictures (in the Animal task) was not counteracted by an opposing force inhibiting response times to the minor chord in the category of musical instruments.

One peculiarity in the present set of data remains. Observe that in Figures 5.1 and 5.2 only in the Instrument task, only for the pictures of Instruments, trials with sound are (at least numerically) faster than trials without sound (this numerical trend is marginally significant in blocks 3 & 4, $t(23) = 1.996$, $SEM = 8.48$, $p = .058$). We hesitate to attribute a role to this phenomenon that is relevant for the present study, although we would like to argue that the inhibition of sound that we observe in all other (sub)sets of data is contrary to the idea common in human–factors research that, in general, the addition of sound is beneficial for performance.

In sum, the present study shows that homogeneity of a category can be a confounding factor in studies using or investigating the animate versus inanimate distinction. The confound should be carefully controlled as it may require different interpretations of experimental results like, for instance, in the study by Brousseau and Buchanan (2004). At least, the attribution of the picture category to the animacy effect by Lemmens et al. (2004) has been shown to be incorrect. The degree of homogeneity also has an impact on the multimodal affective–congruency effect although a conclusion on the precise effects requires further experimentation. More research is of course required to define the notion of homogeneity.

CHAPTER 6

Summary and conclusions

Shall I refuse my dinner because I do not fully understand the process of digestion?

Oliver Heaviside (1850-1925)
English physicist

Abstract

This thesis presented research and a review of literature on the domains of stimulus–response compatibility (SRC), affective–compatibility effects and affective computing, and how the affective connotation of the major and minor modes in Western tonal music can be used to create auditory feedback–signals that can communicate affect to a perceiver. The work started with an emphasis on human–factors research based on relatively simple assumptions, but the focus gradually shifted to answering research questions of a more basic nature because we encountered the limitations of our assumptions and method. Nevertheless, the general theme always was the multimodal affective–congruency effect that was initially observed by Van Esch–Bussemakers (2001, although she did not yet label the effect as such). We used her paradigm to investigate the potential impact of affective–compatibility effects in affective–computing research (Picard, 1997), and exploited that paradigm to explore the cognitive mechanisms involved in determining the affective–congruency effect.

In this chapter we summarize the review that we presented in Chapter 2 and reiterate the results from the studies reported in Chapters 3, 4, and 5, focusing on the reversed congruency effects that we obtained. We also address some of the implications of our findings for human–factors research. In addition, we discuss the experimental method that we used and describe the contours of a tentative model of hot and cold cognition that encapsulates our findings and our reasoning on the multimodal affective–congruency effect.

6.1 Summary of the chapters

Following the introductory chapter (Ch. 1), this thesis continued with a review of literature on SRC phenomena (Ch. 2) that we considered to be relevant topics for research into affective computing. Affective computing is a relatively recent addition to human–factors research that tries to improve human–computer interaction by incorporating aspects of the well–developed human–human interaction, in this case by embedding capabilities in interfaces to recognize and communicate emotion (see Picard, 1997). Three studies followed, reported in Chapters 3, 4, and 5, that all focused on the multimodal affective–congruency effect. Most of this work was carried out in the later stages of the project; some of it in parallel. We chose this particular order of presentation because it follows the chronology of our work and reflects the evolution of the paradigm and method that we used.

Chapter 2

After the brief introduction in Chapter 1 of the main themes (stimulus–response compatibility, multimodal information–processing, and affective computing), Chapter 2 further elaborated them. The main goal of the chapter was to argue that research in affective computing could benefit from the findings from typical stimulus–response compatibility (SRC) literature by showing that the existence of affective–compatibility effects is indicative of an interaction between hot and cold cognition. Furthermore, the affective Simon task was presented as a special case of SRC paradigms to investigate affective–compatibility effects in line with the taxonomy of SRC paradigms that was proposed by Kornblum et al. (1990). The major/minor distinction in Western tonal music, referring to a positive interpretation of music played in major mode compared to a more negative interpretation of music written in minor mode, was introduced as one way to attach affective valence to affectively–neutral stimuli that is specifically suitable for affective (human–factors) research employing multimodal stimuli. We concluded that SRC paradigms, in the affective domain, are a viable alternative to the cumbersome although direct measurements of the autonomous nervous system's responses. We proposed that the major/minor distinction can be used as an additional transformation of affectively–neutral (families of) earcons (short musical sounds to aid in, for instance, navigation; see Blattner et al., 1989) to generate variations that are capable of communicating affectively–charged information.

Chapter 3

Chapter 3 presented a study in which the transformation to major or minor mode was used to show that it can indeed generate earcons that are capable of communicating affect. In an animate/inanimate picture–categorization task, earcons in major and minor mode were simultaneously presented with visual targets. Because participants had to execute positive or negative responses, a dimensional overlap between the irrelevant stimulus (the earcons) and the response was created. We observed that participants were indeed slower to respond to specific combinations of affective valence of stimulus and response. Surprisingly, however, longer response–time latencies were observed for the more natural congruent condition (i.e., positive responses to trials with major chords and negative responses to trials with minor chords) whereas the RT's from the incongruent condition were short.

Assuming that major and minor chords do have the affective connotation of positive valence and negative valence, respectively, and that the more natural compatible conditions should normally show improved performance compared to less natural incompatible conditions in SRC tasks, we proposed that participants changed their information–processing strategy in a way that may have influenced how they determined the instructed responses. This change must have incorporated an evaluation with an affective charge opposite to that of the instructed response. For instance, if the supposed natural (and thus faster) combination of positive valence (of the instructed response) and major chord actually is slower than the supposed less natural (and thus slower) combination of positive valence and minor chord, the former combination must have been an incompatible one for the participants' cognitive system.

Tentatively we suggested that our assumption that Yes–responses are associated with positive affective–valence (and v.v. for No–responses) was violated in the Instrument⟶Yes part of the experiment. In the latter part of the experiments participants were (still) instructed to determine their responses by answering the question "Is the picture you see that of an animal?", but by verbal and written examples they were, essentially, requested to give the wrong answer (i.e., to press the No–button for animals and the Yes–button for the instruments). This may have resulted in Yes–responses with negative valence and No–responses with positive valence. We also postulated that, because of their shared musical character, the relationship between the pictures of the musical instruments and the earcons in major and minor mode was stronger than that of the animals and the earcons. We suggested that these two effects may have accumulated into a relatively large reversed affective–congruency effect for the instrument pictures in the Instrument⟶Yes task, that in turn resulted in an overestimation of the overall congruency effect.

So in Chapter 3, we showed that a transformation to major or minor mode indeed generated earcons capable of communicating affect and that the differences between the congruent and incongruent conditions showed that it is important to maintain affective correspondence in affective human–computer interfaces. Our manipulation to switch from a positive response assigned to the animate category to a negative response highlighted that participants often choose response–selection methods other than the one(s) intended by the experimenter.

Chapter 4

Chapter 4 explored the multimodal affective–congruency effect in more detail than the study reported in Chapter 3, using slightly different experimental materials. In a picture–categorization task one group of participants was instructed to execute positive responses to the targets from the animate category whereas another group of participants was instructed to execute negative responses. For the former group we found a significant multimodal affective–congruency effect with shorter RT's for the congruent condition and longer RT's for the incongruent condition. The group executing the negative responses, however, showed a reversed congruency effect: The congruent condition showed longer RT's whereas the incongruent condition showed shorter RT's. The reversed affective–congruency effect was also significant.

The second–order interaction between the response assignments and the congruency manipulation is a prediction typical for the judgemental–tendency model that Klauer and Stern proposed in 1992. We therefore concluded that judgemental tendencies played an important role in the realization of the multimodal affective–congruency effect. We reasoned that the logically–complex instruction that we used for the group of participants that had to press the No–button to pictures from the animate category, triggered a (subconscious) search for a simpler response–selection procedure and proposed that a judgemental tendency as a result of the emphasis on the animate category (see De Houwer et al., 2002) may have provided a simpler translation.

In sum, in Chapter 4 we obtained further evidence of the multimodal affective–congruency effect corroborating the findings in Chapter 3 and we showed that the judgemental–tendency model, in retrospect, predicted our findings. The judgemental tendencies were proposed as a strategy to realize a simpler stimulus–response translation.

Chapter 5

Besides the multimodal affective–congruency effect, in Chapter 3 as well as in Chapter 4, a consistent processing advantage was obtained for the category of animals. Similar processing advantages have been obtained in other studies employing the animate versus inanimate distinction (e.g., Brousseau & Buchanan, 2004; de Groot, 1990; von Studnitz & Green, 2002) and are sometimes reported to have their origin in the importance of the animate category during the evolution of human cognition (Caramazza & Shelton, 1998).

In the third and final experimental study that we report in Chapter 5, we examined the animacy effect more closely. By carefully balancing the experimental design and maintaining equivalent categorical homogeneity across the animate and inanimate objects, we showed that it was not a general processing advantage for the animate category explaining the animacy effect in Chapters 3 and 4 but differences in processing times due to different categorization procedures (exemplar–based categorization vs. hypothesis–based categorization; see Rosch & Lloyd, 1978). The different categorization procedures were the result of differences in class homogeneity between the animate and inanimate categories. We showed that the emphasis on one specific category (that was related to the judgemental tendencies in Chapter 4) did result in a processing advantage and that this was not due to differences in processing times for selection of positive or negative responses.

So in Chapter 5 we showed that it is important to ensure that categories of stimuli are equivalently homogeneous, because failing to do so may introduce artifacts in the (main) experimental findings.

6.2 Conclusions

The general findings in this thesis may be best summarized as follows. The repeated findings of multimodal affective–congruency effects demonstrate the intriguing but relatively unexplored interface between hot and cold cognition and indicate, for the first time to our knowledge (but see Van Esch–Bussemakers, 2001) that affectively–charged information that is either presented auditorily or resulting from the output system generating positive or negative responses, is processed via the semantic system and this affective information spreads across modalities. The (regular) congruency effect, that we found for the group of participants that were instructed to execute positive responses to the pictures of animals, indicates that our assumption of positive affect associated with the auditory flankers in major mode has not been invalidated (and v.v. for the minor mode and negative affect). The specific (experimental) setup used to convey affect between man and machine proved a decisive factor, however, as indicated by the role of judgemental tendencies (Klauer & Stern, 1992) in the realization of the multimodal affective–congruency effect. We will now first deal with the findings that are relevant for human–factors research and then elaborate on some methodological and theoretical aspects of the reported studies.

6.2.1 Important issues for human–factors research

Although we did not develop a new human–computer interface or explicitly formulated guidelines for such a new interface, human–factors research has received a considerable amount of attention in this thesis. In particular, we elaborated on the domain of affective computing and discussed the potential usability of stimulus–response compatibility (SRC) paradigms for research in that area. In this section we focus on some of the results of the studies reported in this thesis that we consider to be important for future human–factors research.

Based on the importance of affect and emotion in human–human interaction, affective computing is a worthwhile human–factors research effort to investigate the usability of affect to improve the quality of human–computer interaction (Picard, 1997). Our repeated findings of affective–congruency effects provide evidence to suggest that the interaction between hot and cold cognition is important for task performance. It is therefore relevant to maintain affective correspondence between information presented in different modalities to ensure that performance in an affective human–computer interface stays at an optimal level. The affective–congruency effects also show that it is possible to investigate performance in tasks employing affectively–charged stimuli and/or responses using relatively simple behavioral measurements in addition to the more complex (neuro)physiological measurements of EEG, EMG, the galvanic skin–response, and blood–volume pressure.

The knowledge that affectively–neutral earcons can be made affectively charged using a transformation into major or minor mode can be readily used by sound–designers that need to create such earcons for affective human–computer interfaces. The affectively–charged earcons can be used, for instance, to more easily distinguish between the emoticons expressing positive moods and those expressing negative affect when the emoticons are displayed in the long lists that can be selected in many email or online–presence software–packages.

The (overall) reversed congruency effect that we have obtained in Chapters 3 and for one group of participants in the study reported in Chapter 4 show that sometimes it is difficult to a priori determine the congruency relations in an experiment. Regarding human–factors research, we conclude that is important to interpret human behavior within a realistic context rather than in the context of isolated to devices, stimuli, or stimulus–response relations (e.g., see Stammers, Carey, & Astley, 1990; Strater & Bubb, 2003). The type and complexity of an instruction play an important role in how participants (subconsciously) determine the most efficient response–selection strategy. The chosen translation can be different from the one that the experimenter expects based on his instruction. The alternative response–selection strategy may in turn result in different congruency effects than originally expected (e.g., reversed with slower congruent conditions). This is an important aspect for human–factors research to take into consideration because it indicates that behavior and responses should be measured not only in the reduced reality of the controlled laboratory situations but also in real–life situations with devices that are operated in everyday activities.

Finally, in line with Van Esch–Bussemakers (2001), we argue that sound is not always the performance benefactor that it is considered to be (e.g. see the auditory accessory effect; Keuss et al., 1990). The earcons (also in major/minor mode) that Van Esch–Bussemakers (2001) employed (see Bussemakers, de Haan, & Lemmens, 1999; Lemmens et al., 2000, 2001) have more often shown performance degradation than performance enhancement, when measured, for instance, in terms of response–time latencies. This is also the case for the experiments reported in this dissertation: the data consistently show slower responses for nearly all conditions with sound than for the trials without sound extracted from the same condition(s)[1]. We conclude that our participants always process the auditory flankers to a certain extent and that they can neither ignore the sounds nor shut them out completely, because humans lack the physical ability to close the ear canal in the same way that we can close our eyelids to prevent distal stimuli from entering our visual system. Perhaps the level of importance that is attributed in the instruction to the auditory flankers is an important factor determining the degree of the performance degradation or enhancement. Task–performance levels follow an inverted–U curve depending on acoustic stimulation. We may have failed to realize a setting in which optimal performance is ensured. Considerate use of sound in human–computer interfaces, especially in time–critical tasks, is therefore well advised. Despite that multimodal perception is a fact of life for humans, its underlying processes are not always transparent nor is multimodal perception always the fastest, simplest, or most accurate mode of perception.

6.2.2 Remarks on the experimental paradigm

In this section we will comment on the paradigm that we have used in the studies reported in this thesis to try and address critical remarks that a reader may have. First, we discuss the stimuli that we employed and then report on the experimental design. Finally, we discuss the paradigm itself.

---

[1] Two exceptions to this rule can be found in the righthand panel of Figure 4.2 and 5.1.

Materials

A potential concern with respect to the experimental materials that we employed is that there were relatively few variations in the visual targets and no variation in the auditory flankers. In the appendices to Chapter 4 we already confirmed our assumptions on the affective valence of the auditory stimuli (Appendix 4.A, p. 63) and showed that a potential affective charge of the visual targets was not involved in our findings and could not account for other compatibility relations (see Appendix 4.B, p. 67).

Nevertheless, despite the finding that our auditory flankers in major or minor mode were rated differently on affective charge, the affective charge itself may have been modest due to the fact that we employed artificially generated auditory flankers. The modesty of the affective charge of the auditory flankers in major or minor mode could then, partially, account for the difficulties we encountered in finding the multimodal affective–congruency effect. For instance, according to the dual–route models, particularly, that are commonly used to explain many findings in SRC research (see de Jong et al., 1994), compatibility effects stem from quicker processing of the task–irrelevant stimulus feature along a direct priming–like route. The quicker processing results in the pre–activation of a response code based on the irrelevant stimulus that may or may not overlap with the response code as a result of the (slower) processing of the task–relevant stimulus feature (Ridderinkhof, 2002). In our case, the possibly diminished affective charge of the auditory flankers may result in slower, or less efficient processing along the direct priming route, in turn resulting in smaller affective–congruency effects. Crowder (1985a), however, used chords consisting of sine waves that sounded even more artificial than the ones used in our experiments. Nevertheless, he observed a clear separation between major and minor chords with the former being associated with positive affective valence and the latter with negative valence (Crowder, 1985b). Moreover, using recordings of a concert piano, on which the same major and minor chords were played, did not result in a stronger affective–congruency effect (Lemmens & de Haan, 2001). Therefore, we tentatively conclude that it is unlikely that the affective charge was insufficiently pronounced within our auditory flankers.

One tactic to investigate whether the strength of the affective charge of auditory flankers is important in realizing the affective–congruency effect is to use flankers with a pronounced affective charge, for instance, the sound of fingernails on a blackboard versus the sound of a purring kitten. These sounds typically have a negative and positive affective valence, respectively. The perception (and appreciation) of these types of sounds, however, is dependent on individual likings. Some people are not sensitive to the fingernails sound and other people have an utter dislike of cats and kittens. Moreover, if a participant is sensitive to the fingernails sound then the appraisal of that sound may well result in distress (Ortony, Clore, & Collins, 1988) or even an avoidance response instead of a negatively valenced appreciation. Of course, other synthetically generated sounds can be conceived of that are perceived as pleasant and unpleasant. We did not use such sounds, however, because of our interest to investigate whether earcons, which are short *musical* fragments that serve as feedback signals in human–computer interfaces (Blattner et al., 1989), can be enhanced with an affective charge.

Finally, we would like to reiterate our findings on the animacy effect. Throughout the studies that are reported in this thesis and elsewhere (e.g., Lemmens et al., 2000, 2001, 2003) we have consistently obtained a processing advantage for the pictures of animals. Initially, we considered this processing advantage to be the animacy effect that other researchers also obtained (e.g., von Studnitz & Green, 2002). The animacy effect may reflect the importance of the animate category that was established during the evolution of human cognition and may even be physically encoded in our brain (Caramazza & Shelton, 1998). However, closer inspection (Ch. 5) of the stimuli used in the study reported in Chapter 4 showed that the emphasis in the instruction on one of the categories was the likely cause of the shorter response–time latencies. We therefore conclude that maintaining equivalent categorical homogeneity is import to prevent artifactual response–time differences from distorting the key experimental manipulation(s). This is, of course, not only the case for the animate/inanimate distinction but for any other set of categories.

A blocked vs. a randomized design

The second issue concerns the blocked design that we used throughout the studies reported in this thesis. Such a design, in which the congruency relationship is kept constant in a trial block, is quite uncommon in SRC research. Our choice to continue using such a design was motivated by historical reasons as well as a simple choice to further investigate the affective-congruency effects that we obtained using the blocked design, instead of investigating why we did not find congruency effects in a randomized design. We currently entertain two explanations for the absence of congruency effects in randomized designs.

First, it may be that the dimensional overlap between the affective valence of the auditory flankers and the affective property of the responses was less obvious than we had expected and thus sufficiently obscured in the randomized design to escape reliable detection. A blocked design, however, provides more opportunities to detect the dimensional overlap because, throughout the duration of a congruent or incongruent trial block, the relationship between flanker and picture category, and thus responses, is kept constant. In other words, returning to the argument of the possibly diminished strength of the affective charge, the blocked design ensured that the affective charge of the auditory flankers could be picked up even more reliably, which in turn resulted in larger affective–congruency effects. Note that, because of the predictive role of the auditory flankers, we always intermixed an equal number of trials with and without sound within a trial block, to try and prevent participants from exploiting the 100% valid relationship between flanker and target. Also note that participants were consistently slower when responding to trials accompanied by auditory flankers, so it does not seem to be the case that they were able to ignore the flankers completely. The rating study reported in Appendix 4.A shows that flankers in major mode were rated as positively charged whereas the flankers in minor mode were rated as negatively charged, so, regarding the role of the auditory flankers, it seems unlikely that their affective charge could not be exploited by the cognitive system.

We position our second explanation on the level of the stimulus–response sequence(s) by observing that the duration of the auditory flankers was much longer relative to the mean overall response–time latency of around 450 ms. In the studies reported in this thesis we have used auditory flankers with a duration of 2500 ms and other studies have used flankers of 1250 ms (e.g. see Lemmens et al., 2001). We propose that the extended duration of the flankers may have disturbed the normal experimental sequence of stimulus, response execution, (possible) feedback, and inter–trial interval for the auditory stimuli specifically, by blurring the border between an identity as auditory flanker next to a (visual) stimulus and an identity as auditory feedback signal. As a result, participants may have interpreted the auditory flankers differently. Some participants (perhaps those that are auditorily dominated, Giard & Peronnet, 1999) may have interpreted the auditory flankers as such, whereas other participants might have used them as (awkward) feedback signals. Of course, this line of reasoning warrants further study. Using a randomized design, such a study can exploit briefly presented auditory flankers of around 300 ms, that are known to still have the conventional affective connotation (Crowder, 1985a, 1985b).

The experimental paradigm overall

One could remark that more variation in our experimental paradigm could perhaps have clarified our sometimes ambiguous results. In our view, however, we have shown that the materials that we used to build our paradigm adhered to our assumptions and therefore could not account for our findings by the introduction of an artifact. One part of understanding results is to understand the quirks of a certain paradigm; the latter can be done only by carefully manipulating single aspects of the same paradigm to verify the impact of that aspect on the results.

Furthermore, the constancy of the design does not necessarily deteriorate the value of our findings. For instance, despite that participants have ample opportunity to try and learn to ignore the auditory flankers (which leads to performance benefits given the shorter response–time latencies to trials without sound), our findings show a general slowing of RT's in every experiment reported here (and elsewhere, e.g., Lemmens et al., 2000, 2001). From these findings we conclude that ignoring the auditory flankers was not part of an energy–conservation strategy to reduce the distracting factors in a task–set to a minimum. In other words, participants seem to have been continuously aware that the auditory flankers were or were not contingent with the responses.

The multimodal affective–congruency effect itself serves as another example, here. The finding that it is realized only in blocked designs hints at the involvement of strategies or effects of learning or conditioning, instead of the natural overlap between stimulus– or response–features that are common in SRC research. Nevertheless, it still is the overlap between the affective charge of the flankers and the responses that results in two qualitatively different conditions for which (statistically) different response–time latencies are observed. Therefore the multimodal affective–congruency effect shows that it is important to maintain affective correspondence, especially in situations where correspondence relations, that are similar to ours, repeatedly occur.

6.2.3 The involvement of judgemental tendencies

After having obtained a multimodal affective–congruency effect in the study reported in Chapter 3 using a variation of the design employed by Van Esch–Bussemakers (2001, Exp. 3, Ch. 2), we tried to determine the cognitive basis of the congruency effect in Chapter 4. We found that the response–assignment manipulation to ensure that positive as well as negative responses were assigned to both categories of pictures, interacted with our predetermined conception of affective congruency. That is, the participants that were instructed to execute a No–response to pictures from the animate category were faster doing so in the condition that was a priori determined to be incongruent and they were slower in the a priori determined congruent condition.

Klauer and Stern (1992) proposed the judgemental–tendency model to explain how attitudes guide memory–based judgements. The second–order interaction between response assignment and congruency that we also observed in our study is typical of the judgemental–tendency model. According to the model, first a response tendency to confirm or reject the automatically–extracted affective relationship of, for instance, prime and target is constructed. This tendency is then, in a controlled process, compared to the instructed response that can be compatible or incompatible with the response tendency. For example, for a prime–target pair like SUMMER–HONEST, a response tendency to confirm this relation is created. If participants are instructed to execute positive responses to word targets (e.g., in a lexical–decision task), then the response tendency is compatible with the instructed (positive) response. However, if the instructed response is a negatively–valenced response, then the response tendency and the instructed response have an incompatible relationship and slower responses are expected despite the fact that the prime–target pair is affectively congruent. The response tendency is also referred to as a judgemental tendency (De Houwer et al., 2002; Wentura, 2000).

The involvement of judgemental tendencies in the realization of the multimodal affective–congruency effect is not unreasonable given that, in our view, the affective–congruency effect represents an interaction between hot and cold cognition. The judgemental–tendency model (Klauer & Stern, 1992) is an elegant formulation of this interaction incorporating aspects of hot as well as cold cognition. In this respect, the model may be further exploited in the domain of cognitive control, specifically in action– and error–monitoring research investigating the error–related negativity (ERN). The ERN is an ERP component thought to be related to the detection of errors during stimulus and response processing (for a recent review see de Bruijn, 2004).

Regarding the interpretation of the ERN, two competing theories are proposed: the reinforcement–learning theory (Holroyd & Coles, 2002) versus the conflict theory (Botvinick, Braver, Barch, Carter, & Cohen, 2001). However, a third proposal exists that relates the error–related negativity to affectively–motivated evaluations of actions and errors. The affective theory of the ERN (Gehring & Willoughby, 2002; Pailing, Segalowitz, Dywan, & Davies, 2002) hypothesizes that the ERN reflects the emotional response to an error or the affective or motivational evaluation of that error because the detection of "risks and opportunities is the fundamental function of affective and emotional processing" (Wentura, 2000, p. 467; also see Frijda, 1988; Simon, 1967). This theory has not been investigated extensively mostly because it is difficult to separate the cognitive and affective factors that contribute to the ERN as these are highly dependent (Yeung, 2004). In our view, it may be a worthwhile research effort to investigate whether the judgemental–tendency model can be used as a tool to disentangle the cognitive and affective factors that contribute to the ERN.

6.2.4 Contours of a general model of hot and cold cognition

In experimental psychology, the processing of information related to emotion and affect (also called "hot" cognition, see Abelson, 1963; Hollnagel, 1999) for a long time has been considered to be completely separated from rational decision–making ("cold" cognition; e.g. see Sorrentino & Higgins, 1986a).[2] However, in line with other findings of affective–compatibility effects (e.g., De Houwer & Eelen, 1998; De Houwer & Randell, 2004; Fazio et al., 1986), the multimodal affective–congruency effect and particularly the reversed affective–congruency effect (that we observed for one group of participants of the study reported in Chapter 4), in our view, show that cold cognition can be influenced by hot cognition: The

---

[2] Note that, regarding the model that we outline in this section, the use of the undifferentiated term "hot cognition" may require some elaboration. Theories of emotion processing usually distinguish between primary or basic approach and avoidance emotions and more cortical or cognitively determined emotions like pride and inspiration (Ortony et al., 1988; Van Egmond et al., 2004). We regard the affective valence of our auditory flankers as part of the latter type of emotion.

evaluation of (task–irrelevant) affective information present in the stimulus influences the interpretation or realization of the instructed responses. We suggest that hot cognition influences cold cognition in such a way that the (entire) cognitive system remains in an affectively–consistent state. Therefore, cognition (actively) tries to maintain information–processing strategies that do not result in affectively–inconsistent states.

It does so by (automatically; see Fazio et al., 1986) extracting all affective information present in its surroundings and evaluating the information for affective consistency (Klauer & Stern, 1992; Wentura, 2000). If some of the affective properties are inconsistent with others, cold cognition suffers as performance measures on simple rational decisions show (e.g., RT's or errors; see De Houwer & Eelen, 1998; De Houwer & Randell, 2004; Wentura, 2000). Moreover, if inconsistency is detected, the cognitive system tries to return to a more consistent state by (actively) changing those aspects of information processing that are affectively inconsistent to a state in which they are consistent with the other affective features of information processing. We propose that hot cognition wraps like a blanket around cold cognition and that the radiant heat of hot cognition permeates cold cognition, slowly warming up the latter in an effort to make the entire cognitive system as affectively–consistent as possible.

The blanket model of the interface and interaction between hot and cold cognition is in agreement with the spreading–activation account of affective–compatibility effects that has received a lot of attention (e.g. see De Houwer & Randell, 2004; Spruyt et al., 2004) whether this account is framed in the traditional semantic–network version or in a PDP version (Wentura, 2000). In the former version of the spreading–activation account usually emotion nodes are proposed to mediate the representation of affective valence (Bower, 1981). Semantic nodes that are positively or negatively charged connect to the respective emotion node. In this way, the emotion nodes function much like the language nodes in the BIA model for bilingual language perception (Dijkstra & Van Heuven, 1998): through the connections with the emotion node(s) "affectively–charged activation" spreads to the semantic nodes, and the emotion nodes themselves are activated merely by processing affectively–charged information (Wentura, 2000). According to the blanket model, the emotion nodes are, of course, part of the blanket and, once activated, spread their influence into the colder part(s) of cognition. Their influence, for instance, results in affective–priming effects for prime–target pairs that are semantically unrelated.

The blanket model capitalizes more on the process of verifying the consistency of the affective state of the cognitive system when the judgemental–tendency model, another important model for affective–compatibility effects, is concerned. Again, the processing of stimuli automatically extracts affective information that activates the emotion nodes. But in this case, the influence of the affective information spreads beyond the pre–activation (or priming) of semantic concept nodes. This is indicated by the (multimodal) affective–congruency effect that is due to the dimensional overlap between the auditory flankers and the responses. Following the judgemental–tendency model, the affective features present in the stimulus–response array are compared and verified for affective consistency. In addition, if, according to the blanket model, affective consistency can be better maintained using a different stimulus–response translation than the one instructed, the alternative translation is preferred and elected above the instructed one. It is clear that the heat of the blanket of hot cognition extends to processes of decision making (also see Brave & Nass, 2003; Hoffman, 1986) and response selection (this thesis). Note that the above description of the blanket model only provides the outline of our tentative model.

In conclusion, in this thesis we presented three studies and a theoretical review that show that hot cognition can influence cold rational decision–making, corroborating earlier findings of affective–compatibility effects (e.g., De Houwer & Eelen, 1998). Moreover, affective information does not need to be presented in the same modality to which the target stimulus is delivered because affective information is processed and compared in a cross–modal fashion. The influences of hot cognition on cold cognition sometimes depend on strategies to minimize cognitive expenditure and to strive for affectively–consistent states of cognition. We propose the blanket model of hot and cold cognition that summarizes our findings and lines of reasoning. The present project has also provided insights that are useful for human–factors research on affective computing.

Discovering and then trying to understand the multimodal affective–congruency effect has been a highly complicated but rewarding pursuit.

References

Abelson, R. P. (1963). Computer simulation of "hot" cognition. In S. S. Tomkins & S. Messick (Eds.), *Computer simulation of personality: Frontier of psychological theory* (pp. 277–298). New–York: Wiley.

Allport, D. A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. In H. Heuer & F. Sanders (Eds.), *Perspectives on perception and action.* Hillsdale: Erlbaum.

Andre, A. D., & Wickens, C. D. (1990). *Display–control compatibility in the cockpit: Guidelines for display layout analysis* (Tech. Rep.). Moffett Field, CA: NASA Ames Research Center.

Beckers, T., De Houwer, J., & Eelen, P. (2002). Automatic integration of non–perceptual action effect features: The case of the associative affective Simon effect. *Psychological Research*, *66*, 166–173.

Bertelson, P. (1999). Ventriloquism: A case of crossmodal perceptual grouping. In G. Aschersleben, T. Bachmann, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 347–362). Amsterdam: Elsevier.

Bertelson, P., Vroomen, J., de Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics*, *62*(2), 321–332.

Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. *Human–Computer Interaction*, *4*, 11–44.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652.

Bower, G. H. (1981). Mood and memory. *The American Psychologist*, *36*, 129–148.

Brave, S., & Nass, C. (2003). Emotion in human–computer interaction. In J. A. Jacko & A. Sears (Eds.), *Handbook of human–computer interaction: Fundamentals, evolving technologies and emerging applications* (pp. 81–96). Hillsdale, NJ: Lawrence Erlbaum Associates.

Brewster, S. A., Wright, P. C., & Edwards, A. D. N. (1994). An evaluation of earcons for use in auditory human–computer interfaces. In S. Ashlund, K. Mullet, A. Henderson, E. Hollnagel, & T. White (Eds.), *Proceedings of interCHI'93* (pp. 222–227). Amsterdam: ACM Press.

Brousseau, G., & Buchanan, L. (2004). Semantic category effect and emotional valence in female university students. *Brain and Language*, *90*, 241–248.

Brown, C. L. (1988). *Human–computer interface design guidelines*. Westport, CT: Ablex Publishing.

Burke, M. W., Gilson, R. D., & Jagacinsky, R. J. (1980). Multi–modal information processing for visual workload relief. *Ergonomics*, *23*, 961–975.

Bussemakers, M. P., de Haan, A., & Lemmens, P. M. C. (1999). The effect of auditory accessory stimuli on picture categorization; implications for interface design. In H. Bullinger & J. Ziegler (Eds.), *Human–computer interaction: Ergonomics and user interfaces, volume 1 of the proceedings of the 8th international conference on human–computer interaction* (Vol. 1, pp. 436–440). London: Lawrence Erlbaum Associates.

Calvert, G. A., Brammer, M. J., & Iversen, S. D. (1998). Crossmodal identification. *Trends in Cognitive Sciences*, *2*, 247–253.

Caramazza, A., & Shelton, J. R. (1998). Domain–specific knowledge systems in the brain: The animate–inanimate distinction. *Journal of Cognitive Neuroscience*, *10*, 1–34.

Collier, W. G., & Hubbard, T. L. (2001). Judgements of happiness, brightness, speed and tempo change of auditory stimuli varying in pitch and tempo. *Psychomusicology*, *17*, 36–55.

Collier, W. G., & Hubbard, T. L. (2004). Musical scales and brightness evaluations: Effects of pitch, direction, and scale mode. *Musicae–Scientiae*, *8*, 151–173.

Crowder, R. G. (1984). Perception of the major/minor distinction: I. Historic and theoretical foundations. *Psychomusicology*, *4*, 3–12.

Crowder, R. G. (1985a). Perception of the major/minor distinction: II. Experimental investigations. *Psychomusicology*, *5*, 3–24.

Crowder, R. G. (1985b). Perception of the major/minor distinction: III. Hedonic, musical, and affective discriminations. *Bulletin of the Psychonomic Society*, *23*, 314–316.

Crowder, R. G., Reznick, J. S., & Rosenkrantz, S. L. (1991). Perception of the major/minor distinction: V. Preferences among infants. *Bulletin of the Psychonomic Society*, *29*, 187–188.

de Bruijn, E. R. A. (2004). *Action monitoring in normal and altered states*. Unpublished doctoral dissertation, NICI, Radboud University Nijmegen.

de Groot, A. M. B. (1990). The locus of the associative–priming effect in the mental lexicon. In D. A. Balota, F. d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 101–123). Hillsdale: Lawrence Erlbaum Associates.

De Houwer, J. (1998). The semantic Simon effect. *The Quarterly Journal of Experimental Psychology*, *51A*, 683–688.

De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology*, *50*, 77–85.

De Houwer, J., Crombez, G., Baeyens, F., & Hermans, D. (2001). On the generality of the affective Simon effect. *Cognition and Emotion*, *15*, 189–206.

De Houwer, J., & Eelen, P. (1998). An affective variant of the Simon paradigm. *Cognition & Emotion*, *12*, 45–61.

De Houwer, J., Hermans, D., & Eelen, P. (1998). Affective Simon effects using facial expressions as affective stimuli. *Zeitschrift für Experimentelle Psychologie*, *45*, 88–98.

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

De Houwer, J., & Randell, T. (2004). Robust affective priming effects in a conditional pronunciation task: Evidence for the semantic representation of evaluative information. *Cognition & Emotion*, *18*, 251–264.

de Jong, R., Liang, C., & Lauber, E. (1994). Conditional and unconditional automaticity: A dual–process model of effects of spatial stimulus–response correspondence. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 731–750.

Dijkstra, A., & Van Heuven, W. J. B. (1998). The BIA–model and bilingual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 189–225). Mahwah, NJ: Erlbaum.

Donders, F. C. (1969). Over de snelheid van psychische processen [On the speed of mental processes]. *Acta Psychologica*, *30*, 412–431. (Original work published 1868)

Donk, M., & Theeuwes, J. (2003). Prioritizing selection of new elements: Bottum–up versus top–down control. *Perception & Psychophysics*, *65*, 1231–1241.

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

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*, 162–169.

Van Egmond, R. (2004). Emotional experience of frequency modulated sounds: Implications for the design of alarm sounds. In D. de Waard, K. A. Brookhuis, & C. M. Weikert (Eds.), *Human factors in design* (pp. 1–12). Maastricht, the Netherlands: Shaker Publishing.

Van Egmond, R., Desmet, P., & Van Der Helm, A. (2004). Basic and cognitive emotions in sound perception. In *Proceedings of design emotion conference.* Ankara, Turkey: Design & Emotion Society.

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

Fitts, P. M., & Deininger, R. L. (1954). S–R compatibility: Correspondence among paired elements within a stimulus and response codes. *Journal of Experimental Psychology*, *48*, 483–492.

Fitts, P. M., & Seeger, C. M. (1953). S–R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, *46*, 199–210.

Fort, A., Delpuech, C., Pernier, J., & Giard, M. (2002). Early auditory–visual interactions in human cortex during nonredundant target identification. *Cognitive Brain Research*, *14*, 20–30.

Fraisse, P. (1981). Multisensory aspects of rhythm. In R. D. Walk & H. L. Pick (Eds.), *Intersensory Perception and Sensory–Integration* (pp. 217–248). New–York: Plenum Press.

Frijda, N. H. (1988). The laws of emotion. *The American Psychologist*, *43*, 349–358.

Frijda, N. H. (1994). Varieties of affect: Emotions and episode, moods, and sentiments. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion.* New York: Oxford University Press.

Frith, C., & Wolpert, D. (2003). *The neuroscience of social interaction*. Oxford: University Press.

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience*, *7*, 773–778.

Gaver, W. (1989). The SonicFinder: An interface that uses auditory icons. *Human–Computer Interaction*, *4*, 67–94.

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid procesesing of monetary gains and losses. *Science*, *295*, 2279–2282.

Giard, M. H., & Peronnet, F. (1999). Auditory–visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, *11*, 473–490.

Gopher, D. A., & Sanders, A. F. (1984). S–Oh–R: Oh Stages! Oh Resources! In W. Prinz & D. A. Gopher (Eds.), *Cognition and motor processes* (pp. 231–253). Berlin: Springer Verlag.

Gregory, A. H., Worrall, L., & Sarge, A. (1996). The developmental control of emotional responses to music in young children. *Motivation and Emotion*, *20*, 341–348.

Hasbroucq, T., & Guiard, Y. (1991). Stimulus–response compatibility and the Simon effect: Toward a conceptual clarification. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 246–266.

Healey, J., Picard, R., & Dabek, F. (1998). A new affect–perceiving interface and its application to personalized music selection. In *Proceedings of the 1998 workshop on perceptual user interfaces PUI'98.* San Fransisco, CA: PUI'98.

Hedge, A., & Marsh, N. W. A. (1975). The effect of irrelevant spatial correspondences on two–choice response–time. *Acta Psychologica*, *39*, 427–439.

Heinlein, C. P. (1928). The affective characters of the major and minor modes in music. *Journal of Comparative Psychology*, *8*, 101–142.

Hermans, D., & De Houwer, J. (1994). Affective and subjective familiarity ratings of 740 Dutch words. *Psychologica Belgica*, *34*, 115–139.

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

Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology*, *48*, 246–268.

Hoffman, M. L. (1986). Affect, cognition, and motivation. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition. Foundations of social behavior* (pp. 244–280). New Yorks: The Guilford Press.

Hollnagel, E. (1999). Keep cool: The value of affective computer interfaces in a real world. In H. Bullinger & J. Ziegler (Eds.), *Human–computer interaction: Ergonomics and user interfaces. Proceedings of the 8th international conference on human–computer interaction* (pp. 676–680). London: Lawrence Erlbaum Associates.

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error–related negativity. *Psychological Review*, *109*, 679–709.

Hommel, B. (1997). Interactions between stimulus–stimulus congruence and stimulus–response compatibility. *Psychological Research*, *59*, 248–260.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*, 849–937.

Howard, D., Rosen, S., & Broad, V. (1992). Major/minor triad identification and discrimination by musically trained and untrained listeners. *Music Perception*, *10*, 205–220.

Howard, I. P., & Templeton, W. B. (1966). *Human spatial orientation*. Oxford: John Wiley and Sons.

von Helmholtz, H. L. F. (1954). *On the sensation of tone*. New York: Dover. (Original work published 1885)

Kantowitz, B. H., & Campbell, J. L. (1996). Human factors in transportation. In M. Mouloua & R. Parasuraman (Eds.), *Automation and human performance: Theory and applications* (pp. 117–136). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kastner, M. P., & Crowder, R. G. (1990). Perception of the major/minor distinction: IV. Emotional connotations in young children. *Music Perception*, *8*, 189–202.

Keuss, P. J. G., van der Zee, F., & van den Bree, M. B. M. (1990). Auditory accessory effects on visual processing. *Acta Psychologica*, *75*, 41–54.

King, A. J., Schnupp, J. W. H., & Doubell, T. P. (2001). The shape of ears to come: Dynamic coding of auditory space. *Trends in Cognitive Sciences*, *5*, 261–270.

Klauer, K. C., & Musch, J. (2001). Does sunshine prime loyal? Affective priming in the naming task. *The Quarterly Journal of Experimental Psychology*, *54A*, 727–751.

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

Koch, I., & Kunde, W. (2002). Verbal response–effect compatibility. *Memory & Cognition*, *30*, 1297–1303.

Kornblum, S. (1994). The way irrelevant dimensions are processed depends on what they overlap with: The case of Stroop– and Simon–like stimuli. *Psychological Research*, *56*, 130–135.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus–response compatibility — A model and taxonomy. *Psychological Review*, *97*, 253–270.

Kornblum, S., & Lee, J. (1995). Stimulus–response compatibility with relevant and irrelevant stimulus dimensions that do and do not overlap with the response. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(4), 855–875.

Kornblum, S., & Stevens, G. T. (2002). Sequential effects of dimensional overlap: Findings and issues. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action, attention and performance XIX* (pp. 11–54). Oxford: Oxford University Press.

Kornblum, S., Stevens, G. T., & Requin, J. (1999). The effects of irrelevant stimuli: 1. The time course of stimulus–stimulus and stimulus–response consistency effects with Stroop–like stimuli, Simon–like tasks, and their factorial combinations. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(3), 688–714.

Kunde, W. (2003). Temporal response–effect compatibility. *Psychological Research*, *67*, 153–159.

Kunde, W., Hoffmann, J., & Zellmann, P. (2002). The impact of anticipated action effects on action planning. *Acta Psychologica*, *109*, 137–155.

Laxar, K., & Olson, G. M. (1978). Human information processing in navigational displays. *Journal of Experimental Psychology: Applied*, *63*, 734–740.

Lemmens, P. M. C., Bussemakers, M. P., & de Haan, A. (2000). The effects of earcons on reaction times and error–rates in a dual task vs. a single task experiment. In P. R. Cook (Ed.), *Proceedings of the international conference on auditory display* (pp. 177–183). Atlanta, GA: International Community for Auditory Display.

Lemmens, P. M. C., Bussemakers, M. P., & de Haan, A. (2001). Effects of auditory icons and earcons on visual categorization: The bigger picture. In J. Hiipakka, N. Zacharov, & T. Takala (Eds.), *Proceedings of the 7th international conference on auditory display* (pp. 117–125). Helsinki: Laboratory of Acoustics and Audio Signal Processing and the Telecommunications Software and Multimedia Laboratory.

Lemmens, P. M. C., & de Haan, A. (2001). *Artificially generated versus natural recordings of major and minor chords.* (Unpublished raw data)

Lemmens, P. M. C., de Haan, A., & van Galen, G. P. (2003). Do location and content operate independently? In E. Brazil & B. Shinn–Cunningham (Eds.), *Proceedings of the international conference on auditory display* (pp. 59–62). Boston, MA, USA: ICAD.

Lemmens, P. M. C., de Haan, A., & van Galen, G. P. (2004). Multimodal affective congruency effects: The role of judgemental tendencies. *Manuscript submitted for publication*.

Lemmens, P. M. C., de Haan, A., van Galen, G. P., & Meulenbroek, R. G. J. (2004a). Effects of animacy and categorical homogeneity in multimodal affective categorization. *Manuscript submitted for publication*.

Lemmens, P. M. C., de Haan, A., van Galen, G. P., & Meulenbroek, R. G. J. (2004b). Emotionally charged earcons reveal affective congruency effects. *Manuscript submitted for publication*.

Lu, C., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review*, *2*, 174–207.

Mandler, J. M. (1997). Development of categorisation: Perceptual and conceptual categories. In G. Bremner, A. Slater, & G. Butterworth (Eds.), *Infant development: Recent advances.* Hove: Psychology Press.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748.

Meredith, C., Edworthy, J., & Rose, D. (1999). Observational studies of auditory warnings on the intensive care unit. In N. A. Stanton & J. Edworthy (Eds.), *Human factors in auditory warnings.* Aldershot: Ashgate.

Meredith, M. A. (2002). On the neuronal basis for multisensory convergence: A brief overview. *Cognitive Brain Research*, *14*, 31–40.

Moors, A., & De Houwer, J. (2001). Automatic appraisal of motivational valence: Motivational affective priming and Simon effects. *Cognition & Emotion*, *15*, 749–766.

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

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions*. Cambridge: Cambridge University Press.

Pailing, P. E., Segalowitz, S. J., Dywan, J., & Davies, P. L. (2002). Error negativity and response control. *Psychophysiology*, *39*, 198–206.

Pashler, H. E. (1998a). *The psychology of attention*. Cambridge, Massachusetts: The MIT Press.

Pashler, H. E. (1998b). Selective attention. In H. E. Pashler (Ed.), *The psychology of attention.* Cambridge, Massachusetts: The MIT Press.

Piaget, J. (1981). *Intelligence and affectivity: Their relationship during childhood*. New York: Basic Books.

Picard, R. W. (1997). *Affective computing*. Cambridge, MA: The MIT Press.

Pittenger, R. A. (2003). *Affective and perceptual judgments of major and minor musical stimuli.* Unpublished doctoral dissertation, Dartmouth College.

Pramana, E., & Leung, Y. K. (1999). Exploring the learnability of structured earcons. In H. Bullinger & J. Ziegler (Eds.), *Human–computer interaction: Ergonomics and user interfaces. Proceedings of the 8th international conference on human–computer interaction* (pp. 555–559). London: Lawrence Erlbaum Associates.

Proctor, R. W., & Reeve, T. G. (1990). *Stimulus–response compatibility: An integrated perspective*. North–Holland: Elsevier Science Publishers.

Rakison, D. H., & Poulin–Dubois, D. (2004). Developmental origin of the animate–inanimate distinction. *Psychological Bulletin*, *127*, 209–228.

Ridderinkhof, K. R. (2002). Activation and suppression in conflict tasks: Empirical clarification through distributional analyses. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action, attention and performance XIX* (pp. 494–519). Oxford: Oxford University Press.

Ridderinkhof, K. R., van der Molen, M. W., & Bashore, T. R. (1995). Limits on the application of additive factors logic: Violations of stage robustness suggest a dual–process architecture to explain flanker effects on target processing. *Acta Psychologica*, *90*, 29–48.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*, 192–233.

Rosch, E., & Lloyd, B. B. (1978). *Cognition and categorization*. Oxford, England: Lawrence Erlbaum.

Rothermund, K., & Wentura, D. (1998). Ein fairer Test für die aktivationsausbreitungshypothese: Untersuchung affektiver Kongruenzeffekte in der Stroop–Aufgabe. *Zeitschrift für Experimentelle Psychologie*, *45*, 120–135.

Sanders, A. F. (1980). Stage analysis of reaction processes. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (pp. 331–354). Amsterdam: North–Holland.

Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica*, *53*, 61–97.

Sanders, A. F. (1990). Issues and trends in the debate on discrete vs. continuous processing of information. *Acta Psychologica*, *74*, 123–167.

Sanders, A. F. (1998). *Elements of human performance: Reaction processes and attention in human skill*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Scheirer, J., Fernandez, R., Klein, J., & Picard, R. W. (2002). Frustrating the user on purpose: A step toward building an affective computer. *Interacting with Computers*, *14*, 92–118.

Schriefers, H., & Meyer, A. S. (1990). Experimental note: Cross–modal, visual–auditory picture–word interference. *Bulletin of the Psychonomic Society*, *28*, 418–420.

Shannon, C. E., & Weaver, W. (1969). *The mathematical theory of communication*. Chicago: University of Illinois Press.

Simon, H. A. (1967). Motivational and emotional controls of cognition. *Psychological Review*, *74*, 29–39.

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology*, *81*, 174–176.

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In N. Cowan & A. Barron (Eds.), *Stimulus–Response Compatibility* (pp. 31–86). North–Holland: Elsevier Science Publishers.

Simon, J. R., & Craft, J. L. (1971). Communicating directional information with an auditory display. *Journal of Applied Psychology*, *55*(3), 241–243.

Simon, J. R., Hinrichs, J. V., & Craft, J. L. (1970). Auditory S–R compatibility: Reaction time as a function of ear–hand correspondence and ear–response–location correspondence. *Journal of Experimental Psychology*, *86*, 97–102.

Simon, J. R., & Rudell, A. P. (1967). Auditory S–R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, *51*, 300–304.

Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? preference for auditory modality in young children. *Child Development*, *74*, 822–833.

Sorrentino, R. M., & Higgins, E. T. (1986a). *Handbook of motivation and cognition. Foundations of social behavior*. New Yorks: The Guilford Press.

Sorrentino, R. M., & Higgins, E. T. (1986b). Motivation and cognition: Warming up to synergism. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition. Foundations of social behavior* (pp. 3–19). New Yorks: The Guilford Press.

Spruyt, A., Hermans, D., Pandelaere, M., De Houwer, J., & Eelen, P. (2004). On the replicability of the affective priming effect in the pronunciation task. *Experimental Psychology*, *51*, 109–115.

Stammers, R. B., Carey, M. S., & Astley, J. A. (1990). Task analysis. In E. N. Corlett & J. R. Wilson (Eds.), *Evaluation of human work: A practical ergonomics methodology* (pp. 134–160). Philadelphia, PA: Taylor and Francis.

Stein, B. E., & Meredith, M. A. (1993). *The Merging of the Senses*. Cambridge, MA: MIT Press.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, *30*, 276–315.

Stevens, G. T., & Kornblum, S. (2000). *Goals and dimensional overlap: The effects of irrelevant response dimensions.* Kloster Irsee, Germany.

Strater, O., & Bubb, H. (2003). Design of systems in settings with remote access to cognitive performance. In E. Hollnagel (Ed.), *Handbook of cognitive task design* (pp. 333–356). Mahwah, NJ: Lawrence Erlbaum Associates.

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

Theeuwes, J. (1994). Stimulus–driven capture and attentional set: Selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 799–806.

Theeuwes, J., Godijn, R., & Pratt, J. (2004). A new estimation of the duration of attentional dwell time. *Psychonomic Bulletin & Review*, *11*, 60–64.

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our eyes do not allways go where we want them to go: Capture of the eyes by new objects. *Psychological Science*, *9*, 379–385.

Tipples, J. (2001). A conceptual replication and extension of the affective Simon effect. *Cognition & Emotion*, *15*, 705–710.

van de Ven, J. G. M. (2002). *Getting a grip on mental workload*. Unpublished doctoral dissertation, NICI, University of Nijmegen.

Van Esch–Bussemakers, M. P. (2001). *Do I hear what I see? The influence of auditory icons and earcons on multimodal integration in visual categorization*. Unpublished doctoral dissertation, NICI, University of Nijmegen.

van Galen, G. P., Liesker, H., & de Haan, A. (2004). Enhanced muscular reactivity to stress in work–related upper extremity disorders, a neuromotor noise perspective. In *Proceedings of the european society of biomechanics (esb) congress 2004.* 's Hertogenbosch, NL: ESB.

van Schie, H. T., Mars, R. B., Coles, M. G., & Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience*, *7*, 549–554.

von Studnitz, R. W., & Green, D. W. (2002). The cost of switching language in a semantic categorization task. *Bilingualism: Language and Cognition*, *5*, 241–251.

Voß, A., Rothermund, K., & Wentura, D. (2003). Estimating the valence of single stimuli: A new variant of the affective Simon task. *Experimental Psychology*, *50*, 86–96.

Vroomen, J., & de Gelder, B. (2000). Sound enhances visual perception: Cross–modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1583–1590.

Vu, K.-P. L., & Proctor, R. W. (2003). Naïve and experienced judgments of stimulus–response compatibility: Implications for interface design. *Ergonomics*, *46*, 169–187.

Walker, B. N. (2002). Magnitude estimation of conceptual data dimensions for use in sonification. *Journal of Experimental Psychology: Applied*, *8*, 211–221.

Walker, B. N., & Ehrenstein, A. (2000). Pitch and pitch change interact in auditory displays. *Journal of Experimental Psychology: Applied*, *6*(1), 15–30.

Ward, R. D., & Marsden, P. H. (2003). Physiological responses to different web page designs. *International Journal of Human–Computer Studies*, *59*, 199–212.

Welch, R. B., DuttonHurt, L. D., & Warren, D. H. (1986). Contributions of audition and vision to temporal rate perception. *Perceptional Psychophysics*, *39*, 294–300.

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

Worringham, C. J., & Beringer, D. B. (1998). Directional stimulus–response compatibility: A test of three alternative principles. *Ergonomics*, *41*, 864–880.

Yeung, N. (2004). Relating cognitive and affective theories of the error–related negativity. In M. Ullsperger & M. Falkenstein (Eds.), *Errors, conflicts, and the brain. current opinions on performance monitoring* (pp. 63–70). Leipzig: MPI of Cognitive Neuroscience.

Zajonc, R. B. (1980). Feeling and thinking; preferences need no inferences. *American Psychologist*, *35*, 151–175.

Author Index

Nederlandstalige samenvatting

Emotie en het verwerken van informatie met een emotionele lading is deel van ons dagelijks leven. Soms is emotie het onderwerp van gesprek; veel vaker, echter, beïnvloedt emotioneel geladen informatie ons rationele gedrag op manieren waarvan we ons niet bewust zijn. Zo is bijvoorbeeld bekend dat een positieve stemming resulteert in flexibelere en creatievere strategieën om puzzels of problemen op te lossen, terwijl een negatieve stemming geassocieerd blijkt met doelgericht en geconcentreerd probleemoplossingsgedrag (Brave & Nass, 2003). De term 'hot cognition' (Abelson, 1963; Hollnagel, 1999) wordt in dit kader vaak gebruikt om naar het verwerken van emotionele ofwel affectieve informatie te verwijzen, terwijl de term 'cold cognition' gebruikt wordt om te refereren naar rationele beslissingen. Onderzoek naar de rol van affectieve informatieverwerking in menselijke cognitie kan resulteren in kennis die gebruikt kan worden om de mens–machine interactie te verbeteren.

Recentelijk formuleerde Picard (1997) een nieuw onderzoeksgebied binnen de cognitieve ergonomie dat zich specifiek richt op de emotionele aspecten van mens–machine interactie. Zij noemde dit gebied 'affective computing'. Een deel van het fundamentele onderzoek naar affectieve informatieverwerking maakt gebruik van stimulus–respons compatibiliteits paradigma's. In dit proefschrift geven we aandacht aan beide thema's in een serie van experimenten waarin we onderzoeken wat de effecten van geluiden zijn in mens–machine interfaces. De vragen die we proberen te beantwoorden zijn: (a) Is het mogelijk om geluiden te creëren die een affectieve lading dragen en (b) kunnen we de effecten van deze lading meten zonder gebruik te maken van metingen van fysiologische reacties? (c) Kunnen deze affectieve geluiden zinvol ingezet worden in een mens–machine interacties? En, tot slot, (d) kunnen we aangeven hoe de affectieve stimuli de informatieverwerking beïnvloeden? Deze thema's en onderzoeksvragen worden in het eerste hoofdstuk van dit proefschrift nader uiteengezet.

De zes hoofdstukken in dit proefschrift richten zich, achtereenvolgens, op het belang van een beter begrip van de raakvlakken tussen 'hot' en 'cold cognition' voor de cognitieve ergonomie (hoofdstukken 1 en 2). Daarna proberen we aan te tonen dat ook in multimodale (d.w.z. audiovisuele) contexten, affectieve informatie invloed kan uitoefenen op rationele beslissingen (hoofdstukken 3 en 4). Vervolgens onderzoeken we hoe het uitwisselen van informatie tussen 'hot' en 'cold cognition' plaatsvindt (hoofdstuk 4) en evalueren we onze onderzoeksmethode (hoofdstuk 5). In hoofdstuk 6, tot slot, geven we een samenvatting van onze bevindingen en schetsen we een tentatief model van de relaties tussen 'hot' en 'cold cognition'.

Hoofdstuk 2 start met een beschrijving van 'affective computing' (Picard, 1997). Daarna introduceren we stimulus–respons compatibiliteits (SRC) onderzoek en beschrijven we hoe de effecten van taak–irrelevante, affectieve informatie gemeten kunnen worden met behulp van compatibiliteitseffecten. We beargumenteren dat compatibiliteitseffecten een waardevolle aanvulling vormen op de directe maar complexe metingen van fysiologische reacties, zoals hartslag en bloeddruk, die vaak binnen 'affective–computing' onderzoek geanalyseerd worden.

Binnen fundamenteel onderzoek naar affectieve informatieverwerking worden vaak compatibiliteitseffecten gebruikt. Het verwerken van affectieve informatie werd lange tijd gezien als compleet gescheiden van de koude, rationele informatieverwerking. In onze review van de literatuur laten we echter zien dat deze stelling onjuist is (zie bijv. Fazio et al., 1986; Sorrentino & Higgins, 1986a). Veel van de taken die gebruikt worden om affectieve informatieverwerking te onderzoeken kunnen eenvoudig in een taxonomie van SRC–taken (Kornblum et al., 1990) geplaatst worden. Als voorbeeld bespreken we de affectieve variant van de Simon–taak (De Houwer & Eelen, 1998; Simon, 1990) die de onderzoeksgebieden van SRC en affectieve informatieverwerking verbindt. Deze taak staat model voor de taak die wij in dit proefschrift gebruiken.

Tot slot bespreken we in dit hoofdstuk het onderscheid in affectieve waarde van de majeur en mineur connotatie in Westerse muziek. We beargumenteren dat de welbekende associatie van een blije, positieve perceptie van muziekwerken geschreven in majeur en een negatieve, bedroefdere interpretatie van muziek in mineur geschikt is om positieve en negatieve lading via de auditieve modaliteit aan te bieden. Het onderscheid tussen majeur en mineur hebben wij in de verdere hoofdstukken gebruikt om stimuli te creëren die we volgens het patroon van een affectieve Simon–taak multimodaal, d.w.z. audiovisueel, aan proefpersonen aanboden.

Hoofdstuk 3 beschrijft een studie waarin we onderzochten of we met behulp van SRC–paradigma's de invloed van de taak–irrelevante, affectief geladen informatie konden ontdekken. We creëerden twee (sets van) akkoorden: C–drieklanken in majeur en dezelfde drieklanken in mineur. Deze geluiden werden gelijktijdig aangeboden met de visuele stimuli die de proefpersonen moesten categoriseren. Omdat de responsen die de proefpersonen moesten uitvoeren ook affectief geladen waren, verwachtten we kortere reactietijden en minder fouten voor de congruente trials (Ja & Majeur; Nee & Mineur) dan voor de incongruente trials (Ja & Mineur; Nee & Majeur).

De resultaten van het experiment dat we in hoofdstuk 3 rapporteerden lieten echter het tegenovergestelde zien. We veronderstelden dat deze resultaten tot stand kwamen door een onverwachte en speciale relatie tussen het muzikale karakter van de plaatjes van de instrumenten en de geluiden die we gebruikten. In het algemeen concludeerden we echter dat de majeur/mineur connotatie de uitvoering van een taak waarin ook affectieve componenten voorkomen, beïnvloedt.

In hoofdstuk 4 onderzoeken we hoe het compatibiliteitseffect uit het 3e hoofdstuk tot stand kwam. Hiertoe voerden we een aantal controlestudies uit en pasten we de methode uit hoofdstuk 3 aan door, bijvoorbeeld, de plaatjes van de muziekinstrumenten te vervangen door (algemenere) plaatjes van niet–levende objecten. We lieten twee verschillende groepen proefpersonen de instructievarianten uitvoeren die in hoofdstuk 3 beschreven staan.

Voor de eerste groep van proefpersonen, die de simpele instructie "Is het een plaatje van een dier?" ontvingen, lieten de resultaten uit het experiment ditmaal een congruentie–effect zien zoals we dat verwachtten. Dat wil zeggen dat de reactietijden op congruente trials korter waren dan die op incongruente trials. De resultaten uit de tweede groep van proefpersonen, die een moeilijkere, logisch–complexe instructie "Is het *geen* plaatje van een dier" ontvingen, waren echter niet volgens onze verwachting.

We vonden dat deze proefpersonen in de congruente conditie, die we als de snellere conditie kenmerkten, langzamer waren dan in de incongruente conditie. Echter, volgens de regels van overlap (Kornblum et al., 1990) tussen affectieve lading van majeur/mineur geluiden en van Ja/Nee–responsen zou deze instructie geen invloed mogen hebben gehad. Immers, majeur en Ja zijn steeds congruent of het Ja–antwoord nu een reactie is op een plaatje van een dier of een plaatje van een niet–

levend object. We verklaren dit resultaat door te verwijzen naar het 'judgemental–tendency' model van Klauer and Stern (1992). Dit model veronderstelt dat er twee stadia zijn in het bepalen wat de invloed is van affectieve informatie in een stimulus. De cruciale voorspelling van dit model is dat er situaties kunnen voorkomen waarin stimulus–respons relaties, die a priori als congruent worden bestempeld, resulteren in gedragsmaten die eigenlijk bij een incongruente stimulus–respons relatie horen. Voor onze tweede groep van proefpersonen betekende dit dat de conditie die wij a priori als congruent kenmerkten, bijv. een plaatje van een schoen (Ja–antwoord) met een majeur akkoord (positieve lading) in werkelijkheid door het cognitief systeem als incongruent gezien werd, bijvoorbeeld omdat het plaatje van de schoen in vergelijking met de referentiecategorie in de vraag (dieren) negatief was en dit incongruent was met het majeur akkoord.

In hoofdstuk 5 proberen we de oorzaak van een eerder geobserveerd stabiel reactietijdvoordeel voor de categorie van dieren te achterhalen. We postuleerden drie verklaringen voor dit reactietijdvoordeel. Het voordeel kon zijn ontstaan (a) door het belang van levende zaken gedurende de evolutie van de mens (Caramazza & Shelton, 1998), of (b) als resultaat van verschillen in de (relatieve) homogeniteit van de plaatjescategorieën (zie b.v. Rosch, 1975; Rosch & Lloyd, 1978), of (c) door de nadruk die we in de instructie legden op de categorie van dieren.

In het experiment gebruikten we weer de plaatjes van muziekinstrumenten als tweede categorie om homogeniteit als oorzaak uit te sluiten. We hanteerden het tussengroepsdesign uit hoofdstuk 4, maar voor de tweede groep van proefpersonen formuleerden we een instructie die ditmaal de categorie van muziekinstrumenten benadrukte. De resultaten toonden eenduidig aan dat het reactietijdvoordeel niet door het onderscheid levend versus niet–levend was ontstaan zoals we eerder concludeerden, maar door de nadruk die in de instructie op de categorie van dieren werd gelegd.

Het zesde en laatste hoofdstuk gebruiken we om de eerdere hoofdstukken samen te vatten en conclusies te formuleren aan de hand van de vragen die we eerder in deze samenvatting stelden. We sluiten het hoofdstuk af met een schets van de contouren van een model van 'hot' en 'cold cognition' dat onze redeneringen en bevindingen samenvat.

We concluderen dat we geslaagd zijn in de opzet om geluiden te creëren die een affectieve lading dragen. Dit blijkt uit de congruentie–effecten die we hebben gevonden. Deze effecten geven tevens aan dat de effecten van een affectieve lading op taakafhankelijke informatieverwerking ook met behulp van SRC–paradigma's gedetecteerd kunnen worden. Deze bevindingen kunnen gebruikt worden in mens–machine interfaces waarin met geluiden affectief geladen informatie gecommuniceerd moet worden. Hierbij is het belangrijk om gelijkheid tussen alle affectieve ladingen na te streven om een optimale mens–machine interactie te waarborgen.

Het fundamentelere deel van onderzoek in dit proefschrift betrof de vraag hoe de invloed van 'hot cognition' op 'cold cognition' tot stand komt en meer specifiek hoe de affectieve lading van de majeur en mineur geluiden invloed heeft op de responskeuze in een categorisatietaak. We tonen aan dat, net als bijvoorbeeld in de affectieve Simon taak (De Houwer & Eelen, 1998), het cognitief systeem in staat is om verschillende bronnen van affectieve informatie te detecteren en te vergelijken op gelijkheid van affectieve lading. Het 'judgemental–tendency' model van Klauer and Stern (1992) verklaart het beste de resultaten die we verkrijgen wanneer proefpersonen een moeilijke, logisch–complexe instructie moeten uitvoeren. Tevens concluderen we dat het effect waarschijnlijk op dezelfde manier tot stand komt als compatibiliteitseffecten in SRC–onderzoek. We komen tot deze conclusie omdat de informatieverwerkingsstappen van het 'judgemental–tendency' model overeen lijken te komen met de beschrijvingen van omgekeerde compatibiliteitseffecten in het 'logical–recoding' model van Hedge and Marsh (1975).

We sluiten het hoofdstuk af met een voorstel voor een model dat onze bevindingen en redeneringen samenvat. In dit model wordt de 'cold cognition' omhuld door de 'hot cognition', waardoor 'cold cognition' opgewarmd kan worden door 'hot cognition'. We veronderstellen daarnaast dat het gehele cognitieve systeem streeft naar een toestand waarin minimale affectieve inconsistentie heerst, uit een strategie tot energiebehoud.

Publications

Journal articles

Lemmens, P. M. C., de Haan, A., van Galen, G. P., & Meulenbroek, R. G.J. (2004). *Stimulus–Response Compatibility and Affective Computing: A Review.* Manuscript submitted for publication to Theoretical Issues in Ergonomics Science.

Lemmens, P. M. C., de Haan, A., van Galen, G. P., & Meulenbroek, R. G.J. (2004). *Emotionally Charged Earcons Reveal Affective Congruency Effects.* Manuscript submitted for publication to Ergonomics.

Lemmens, P. M. C., de Haan, A., & van Galen, G. P. (2004) *Multimodal Affective Congruency Effects: The Role of Judgemental Tendencies.* Manuscript submitted for publication to Acta Psychologica.

Lemmens, P. M. C., de Haan, A., van Galen, G. P., & Meulenbroek, R. G. J. (2004) *Effects of Animacy and Categorical Homogeneity in Multimodal Affective Categorization.* Manuscript submitted for publication to Psychological Research.

Conference proceedings

Lemmens, P. M. C. (2005). Using the Major and Minor Mode to Create Affectively–Charged Earcons. In E. Brazil (Ed.), *Proceedings of ICAD 05-Eleventh Meeting of the International Conference on Auditory Display* (pp. 205–211). Limerick University, Ireland.

Lemmens, P. M. C., de Haan, A., & van Galen, G. P. (2003). Do Location and Content Operate Independently? In Brazil, E. & Shinn-Cunningham, B. (Eds.), *Proceedings of the International Conference on Auditory Display* (pp. 59–62). Boston University (MA), USA.

Lemmens, P. M. C., Bussemakers, M. P., & de Haan, A. (2001). Effects of Auditory Icons and Earcons on Visual Categorization: The Bigger Picture. In J. Hiipakka, N. Zacharov, & T. Takala (Eds.), *Proceedings of the International Conference on Auditory Display* (pp. 117–125). Espoo, Finland.

Lemmens, P. M. C., Bussemakers, M. P., & de Haan, A. (2000). The Effects of Earcons on Reaction Times and Error–rates in a Dual–Task vs. a Single–Task Experiment. In P. R. Cook (Ed.), *Proceedings of the International Conference on Auditory Display* (pp. 177–183). Atlanta, GA, USA.

Bussemakers, M. P., de Haan, A., & Lemmens, P. M. C. (1999). The Effect of Auditory Accessory Stimuli on Picture Categorization; Implications for Interface Design. In H. Bullinger & J. Ziegler (Eds.), *Human–Computer Interaction: Ergonomics and User Interfaces, Volume 1 of the Proceedings of the 8th International Conference on Human–Computer Interaction* (pp. 436–440). London: Lawrence Erlbaum Associates.

Dankwoord

Na bijna vijf jaar noeste arbeid met vele ups en downs is dit boekje dan uiteindelijk toch af. Zonder de hulp van de volgende personen zou dit nooit gelukt zijn.

Gerard van Galen en Ab de Haan wil ik bedanken voor hun begeleiding tijdens dit promotietraject. Beste Gerard, in eerste instantie was je de prof die op een afstandje de progressie van het project volgde. Tijdens onze gesprekken en discussies heb ik altijd genoten van je brede overzicht van de literatuur en de relaties en ideeën die daaruit voortvloeiden. Je hebt Ab en mij altijd goed kunnen ondersteunen in de moeilijke periodes. Ik wil je nadrukkelijk bedanken voor de dagelijkse begeleiding die je me in de laatste maanden van het project gegeven hebt, ondanks dat je toen nog slechts één dag per week op de uni werkte. Ruud Meulenbroek wil ik bedanken voor zijn bereidwilligheid om op het laatste moment in te springen als 2e co–promotor. Je hulp bij het stylistisch maar ook inhoudelijk afronden van de manuscripten was altijd bijzonder constructief en zeer bruikbaar. Je snelheid van werken en oog voor detail worden terecht geroemd onder de (motoriek) junioren!

Beste Ab, voor jou is dit project net zo'n leertraject geweest als voor mij. We hebben vaak heftig gediscussieerd over het verloop en de gemaakte keuzes; gelukkig konden we de discussie altijd met een gezamelijk standpunt afsluiten. Er was ook altijd plaats voor een persoonlijke noot. Ik vind het bijzonder spijtig dat je de laatste maanden van het project niet meer van nabij mee hebt kunnen maken. Hoewel 't vaak zware maanden waren, was het wel één van de interessantste periodes uit het hele traject. Ik hoop dat dit proefschrift voor jou net zo'n grote overwinning is als dat het voor mij is.

Prof. Bekkering, Prof. Ridderinkhof en Dr. van Egmond wil ik bedanken als leden van de manuscriptcommissie voor de door hen gespendeerde tijd en voor het kritische, maar opbouwende commentaar dat zij gaven op het manuscript.

Mijn paranimfen hebben beiden een belangrijke rol gespeeld in mijn wetenschappelijke leven, waarvoor ik hen beide wil bedanken. Samen met Ellen heb ik ruim vijf jaar Cognitiewetenschap gestudeerd en we zijn vrijwel gelijktijdig met een promotieproject begonnen. Buiten het werk om konden we, samen met Karin en Sigrid, met een spelletje Eufraat altijd voor de nodige ontspanning zorgen (hé, waar ken ik die zin van? :)). Met Edita deelde ik gedurende vier jaar een kamer zonder één onvertogen woord. Alhoewel jouw project wetenschappelijk gezien een compleet andere tak van sport was, vertoonde het niettemin veel overeenkomsten met het mijne. Daardoor hebben we veel academische momenten kunnen delen. *Bori se i pobijedi!*

Andere mensen die er altijd klaarstonden voor vakinhoudelijk commentaar waren natuurlijk mijn jonge en oude collega's op het NICI en vrienden op het FCDC en het UMCN. Bokbier, volleybal en VMB gaven onze academische band vaak een persoonlijk tintje! Bedankt voor jullie regelmatige deelname aan mijn experimentjes en testjes wanneer de studenten het weer eens af lieten weten. Bij sommigen van jullie kon ik zelfs buiten kantooruren praten over promoveren en promotieprojecten.

Beppie, Thea en Yvonne bedank ik voor hun administratieve hulp en voor het delen van hun schier onuitputtelijke kennis over administratieve sluipweggetjes en als bronnen van informatie over vanalles (en nog wat). Gerard van Oijen (en via hem zijn collega's van de RTG) stond altijd klaar met raad en daad bij experimentele opstellingen of computerapparatuur in het algemeen. Hubert Voogd heeft de experimentatie–software geschreven. Dennis Pasveer heeft in het laatste half jaar van het project, als bijzonder gedreven en kritische studentassistent, nog veel data voor me verzameld. Bedankt daarvoor! Othmar Schimmel wil ik bedanken voor het analyseren van de oude set van earcons en voor het maken van (in totaal) vijf verschillende sets van nieuwe earcons waarvan ik een groot deel gebruikt heb bij het uitvoeren van de experimenten.

Veel van mijn niet academische *F*riends stonden, zelfs in de nogal luidruchtige omgeving van cafe's en restaurants, altijd klaar met een geïnteresseerd en luisterend oor. De CogW–club is altijd gezellig, of het nu in Jos is of in Warsberg. Marieke van der Linden wil ik nadrukkelijk noemen en bedanken voor haar hulp bij het ontwerpen en maken van de kaft van het proefschrift. *D'r Limburgse groep wil ich bedanke umdat deer nog ummer u plaetsjke in euge agenda en eug huuske hubt veur die Nijmeegse sjtatsluuj!*

Pap en mam, jullie wil ik bedanken voor jullie steun, geduld en vertrouwen in de goede afloop. Tijdens de ruim vijf jaar studeren en bijna even lang werken aan dit promotieproject zullen jullie ongetwijfeld eens gedacht hebben: *"Wat mot van dae jong goan kome"*. Jullie stimulans, vroeger, om vol te houden, veel te lezen en te weten hebben de basis gevormd voor het succesvol afronden van dit proefschrift. Bedankt!

Sigrid, *es litste bis'te der bitste!* Op deze plek worden vele woorden gebruikt om de significantie van de rol van de partner van de promovendus weer te geven. Ik wil me hier beperken tot een perfecte samenvatting in drie woorden: Bedankt voor alles! ... "I love you not only for what you are, but for what I am when I am with you"[3]. Ik hoop nog veel samen met je mee te maken!

---

[3] Elizabeth Barett Browning, 1806–1861

Curriculum vitae

Paul Lemmens werd op 1 oktober 1975 geboren te Heerlen, waarna hij in Schin op Geul, gelegen in het mooie Geuldal in Zuid–Limburg, opgroeide. In mei 1994 behaalde Paul zijn eindexamen VWO bij de scholengemeenschap Sophianum in Gulpen en verhuisde, voor zijn studie aan de Katholieke Universiteit, naar Nijmegen. De propedeuse Informatica, die hij in het najaar van 1995 behaalde, stelde hem in staat verder te gaan met de studie Cognitiewetenschap. Met een stage en afstudeerscriptie op het gebied van de cognitieve ergonomie sloot Paul in februari 2000 deze studie af. Zijn interesse voor onderzoek was geprikkeld en gelukkig kon zijn afstudeerbegeleidster Myra van Esch–Bussemakers een onderzoeksplek realiseren zodat Paul, als haar collega, zich verder kon ontplooien in een promotietraject. Vanaf september 2004 tot en met maart 2005 was Paul aangesteld als docent aan de Radboud Universiteit Nijmegen.

"The only good thesis is a done thesis."

T. Bonebright
at ICAD2005, Limerick.