Implicit Personality Self-Concept

Assessment and Validation

DISSERTATION

zur Erlangung des akademischen Grades Dr. rer. nat.
im Fach Psychologie

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät II
der Humboldt-Universität zu Berlin

vorgelegt von
Dipl.-Psych. Konrad Schnabel
geboren am 30. Juli 1971 in Waiblingen

Präsident der Humbolt-Universität zu Berlin
Prof. Dr. Jürgen Mlynek

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II
Prof. Dr. Elmar Kulke

Gutachter
1. Herr Prof. Dr. Jens B. Asendorpf
2. Herr Prof. Dr. Anthony G. Greenwald
3. Herr Prof. Dr. Karl-Christoph Klauer

Tag der mündlichen Prüfung: 19.04.2004

The purpose of art is not the release of a momentary ejection of adrenaline but rather the gradual, lifelong construction of a state of wonder and serenity.

GLENN GOULD

Content

Index of Tables and Figures

1 Catch *Id* If You Can: The Introduction

"Catch Me If You Can" was the title of a movie that was released last Christmas, 2002. The main characters were a FBI Agent, Carl Hanratty, and a young con artist, Frank W. Abagnale, who were engaged in a cat and mouse game all throughout the film. In the 1960's, Frank W. Abagnale became known as an extremely successful master at forging IDs as well as personal and commercial checks. Since Sigmund Freud's psychoanalytic theory became public at the beginning of the last century, catch *id* if you can attracts the attention of lay and scientific psychology. Freud (1923) considered the id to be the deep, inaccessible part of personality. One may hypothesize about the id impulses that made, for example, someone like Frank W. Abagnale pretend to be someone else at any costs, including the forgery of official documents.

In the last five decades, Social Cognition research has made progress towards finding the tools to identify and measure the "deep, inaccessible" aspects of individuals (e.g., Bless, Fiedler, & Strack, 2003; Greenwald & Banaji, 1995; Wyer & Srull, 1994). Current models conceptualize human behavior and experience as the function of two different systems of information processing, that is, the reflective and the impulsive system (Strack & Deutsch, in press). In the present work, knowledge representations in the reflective and the impulsive system are conceptualized as explicit and implicit representations, respectively. Recently, indirect measures were developed that allow for the assessment of implicit representations. Indirect measures, in contrast to direct questionnaire measures, are chronometric procedures that avoid directly asking the respondents about their judgments. The most influential class of indirect measures used to this date are the Implicit Association Tests (IATs) by Greenwald, McGhee, and Schwartz (1998).

In the present work, I employ indirect measures to assess the *implicit personality self-concept*, that is, implicit representations of one's own personality. In three studies, I explore the following psychometric properties of indirect measures using the traits of shyness, anxiousness, and angriness as examples. First, are indirect measures less fakable than direct measures? Second, what is the convergent validity between the IATs and a new class of indirect measures, the Implicit Association Procedures (IAPs)? Third, do indirect measures increase the prediction of behavior? Fourth, do indirect measures allow for the concurrent assessment of different personality traits?

Before I try to answer these questions I wish to thank the following persons who helped me in my work. First and foremost, I would like to express my gratitude to Elina Yagudayev-Guralnik for stylistic corrections, thoughtful comments, and helpful suggestions concerning my writing. I also would like to thank the lab members of the department of Personality Psychology at Humboldt University, particularly Rainer Banse (now at the University of York), Jaap Denissen, Franz Neyer, and Sarah Teige who shared their theoretical and practical knowledge of psychology during countless collegial chats.

I thank Harald Schneider for technical support, and the following students for their help as experimenters or role play partners: Stefanie Bublitz, Jekatarina Cechini, Andrea Grasse, Susanne Hillenkamp, Vincenzo Kreft, Stephanie Krumnow, Sebastian Kunert, Jana Lüdtke, Dennis Mocigemba, Kristin Müller, Moritz Röhl, Susanne Scheibe, Ulrike Schild, Tanja Schneider, Anja Sussujew, Sarah Teige, Benjamin Uebel, and Anja Weyl.

I am also deeply grateful to Miguel Brendl and Claude Messner for offering the EMA's Turbo Pascal software, and to Boris Egloff and Monika Wiedig for their helpful comments on the emotion inductions that were used in Study 2. My special thanks to the participants whose willingness to engage in the lab experiments made this research possible.

Last but never least, I wish to thank my advisor, Professor Asendorpf, for all his guidance as well as knowledge shared during the preparation of this work. Considering explicit and implicit representations, I think that I learned a great deal.

This research was partly supported by a grant from the German Research Foundation to Jens B. Asendorpf and Rainer Banse (As 59/9)

2 Theory

The first sections of the Theory chapter are devoted to the theoretical conceptualization of the explicit and the implicit personality self-concept. Following this, different indirect measures are discussed, and the Implicit Association Tests are presented in detail. The final section deals with the personality traits of shyness, anxiousness, and angriness that were assessed in the present studies.

2.1 Explicit and Implicit Personality Self-Concept

Individuals process information in two different ways (e.g., Strack & Deutsch, in press). For instance, a person may feel optimistic about her or his life deliberately as a way of positive thinking or automatically due to positive bias. Deliberate thinking and automatic bias, however, differ with respect to how information is processed and how information is made available. In one way, information is processed reflectively, and is accessible through introspection. In the other way, information is processed impulsively, and is accessible only indirectly. The deliberate and the automatic way may be assigned to different systems of information processing, that is, the Reflective and the Impulsive System (Strack & Deutsch, in press). To differentiate between the information representations of both systems at the construct level, representations in the Reflective System are labeled as *explicit representations*, and representations in the Impulsive System as *implicit representations*. Generally, this work deals with the differences and similarities between explicit and implicit representations.

Specifically, the goal of this work is to study explicit and implicit representations of the *personality self-concept*. The personality self-concept may be defined as an associative network containing all of the associations between the concept of self and personality-describing attributes (Asendorpf, Banse, & Mücke, 2002). Personality-describing attributes refer to individual, relatively stable characteristics of the person, yet, do not include pathological attributes (e.g., agoraphobic) as well as cultural or human universals (e.g., German, vertebrate).

This definition of the personality self-concept is in line with Greenwald, Banaji, Rudman, Farnham, Nosek, and Mellot (2002) who argued that information about social objects, social groups, and the self is stored in *Social Knowledge Structures*. Social Knowledge Structures consist of concepts, that is, representations of persons, groups, or attributes, and associations between these concepts. Thus, the representation of one's own

personality, that is, the personality self-concept, is part of the Social Knowledge Structures. Unlike Greenwald et al.'s (2002) self-concept definition, the personality self-concept includes aspects of self-esteem. Thus, associations between the concept of self and attribute concepts containing a positive or negative valence (e.g., agreeable, disagreeable) are also part of the personality self-concept as long as these attributes describe stable, nonpathological interindividual differences. Shyness, anxiousness, and angriness are examples of personality-describing attributes that are not neutral with respect to valence. These attributes or personality traits were studied both as explicit and implicit representations within the personality self-concept.

In brief, explicit and implicit representations are considered as interacting entities that have different ways of transcribing information from the associative store (for a different conceptualization cf. Wilson, Lindsey, & Schooler, 2000). Thus, explicit and implicit representations are not analogous to the distinction between explicit and implicit memory (Fazio & Olson, 2003). Implicit memory refers to learning effects for which individuals lack awareness (e.g., Schacter, 1987). In contrast, implicit representations are not unaware by definition, and differ from explicit representations with respect to how they provide access to the associative store. The associative store contains all of a person's knowledge in terms of elements that are associated by episodic or semantic links (Strack & Deutsch, in press). Social Knowledge Structures (Greenwald et al., 2002) are the part of the associative store that refers to social objects, social groups, or the self. A more specific definition of explicit and implicit representations will be given in the following section, after the Reflective-Impulsive Model from Strack and Deutsch is discussed.

2.2 Reflective and Impulsive Information Processing

Recently, Strack and Deutsch (in press) presented an exemplary *two-systems model* that comprises and expands previous dual-process models (Chaiken & Trope, 1999; Epstein, 1994; Fazio, 1990; Smith & De Coster, 2000; Wilson et al., 2000; Sloman, 1996). The model proposes that perception, thinking, and behavior are functions of two different systems of information processing: the Reflective and the Impulsive System (see *Figure 1*).

In the *Reflective System*, behavior is the result of a *decision process*. The process starts with a perceptual input that is translated into knowledge, that is, a propositional categorization. This induces a reasoning process that leads from a noetic, that is, conscious,

decision to a behavioral decision. For instance, if a young man notices an elderly person in a bus, he generates the proposition “this is an elderly person” by combining the concepts “elderly” and “person” with the relation “is a”. This propositional categorization may be extended with the concepts “standing” and “tired”, and could induce a reasoning process that, for example, it is not good for an elderly person to stand. The reasoning process then leads to the noetic decision that the elderly person had better take a seat. Before the young man makes a behavioral decision, he looks around for a free seat, checks out whether somebody else is ready to offer it, and reflects upon offering his own seat. Finally, he decides to give up his seat and stands up. Intending is what controls his behavior then, until his aim is realized. Intending will eventually stop his behavior when the elderly person gets off at next station.

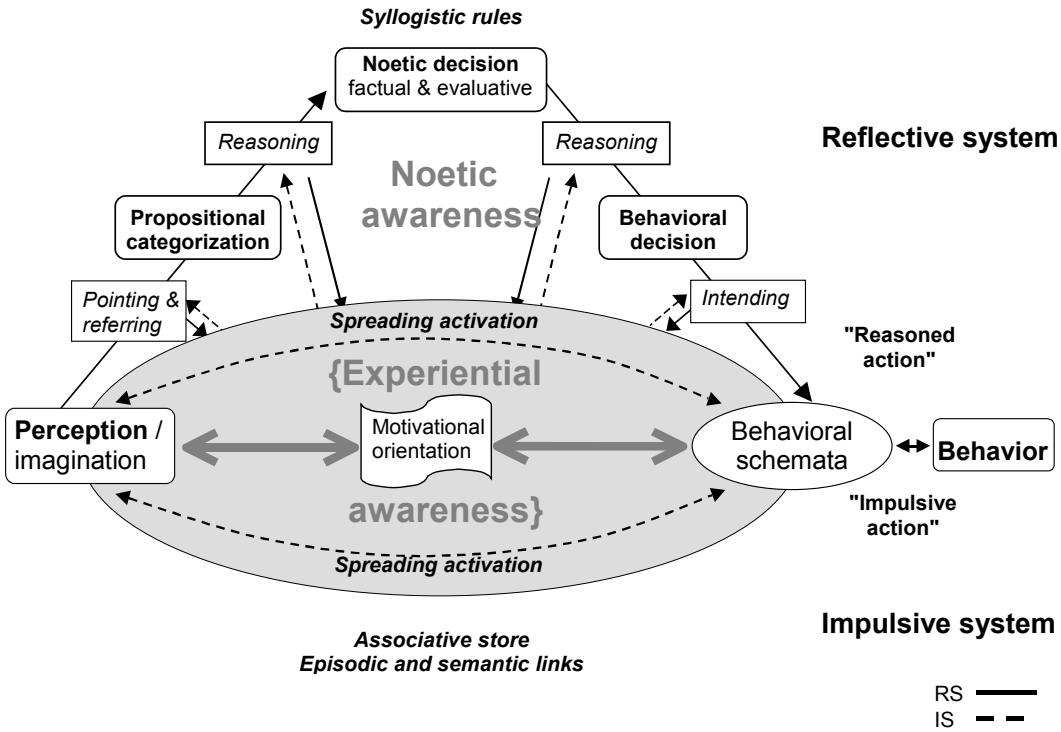

Figure 1. Strack and Deutsch’s (in press) Reflective-Impulsive Model of information processing.

In the *Impulsive* System, behavior is generated by the *spread of activation* from perception and imagination to motor schemata, and by *motivational orientations*. For instance, the young man on the bus may himself move more slowly than he usually does, because the elderly person activated such a stereotype (Bargh, Chen, & Burrows, 1996).

His willingness to offer his seat may be strengthened by approach motivation if the elderly person looks amiable. In contrast, the young man may abstain from offering his seat if the person strengthens his avoidance motivation by looking very unfriendly.

The two systems differ with respect to their structural components, processes, and states. The *structural components* of the Reflective System are concepts that are retrieved from the Impulsive System. These concepts are linked by assigning a truth-value to their relation, whereby the relation is classified as either true or false. The outcome is a propositional categorization. In contrast, the structural components of the Impulsive System are concepts that are associated by episodic and semantic links. These links emerge due to activation in close temporal and spatial contiguity without the assignment of any truth-value. This means that the Impulsive System, in contrast to the Reflective System, is not able to negate information. Whereas the Impulsive System is considered as a long-term storage, the Reflective System has properties of a short-term memory.

Information processing in the Reflective System is a sequence of several decisions that include reasoning and intending. This decision process is flexible, and is able to construct and transform knowledge. Yet, it is slow as it requires intentional resources. In the Impulsive System, information is processed associatively whereby activation spreads using the episodic and semantic links within the associative store. This process is rigid and inflexible but fast.

The *state of awareness* in the Reflective System is described as noetic, that is, it consists of knowledge about the information that is processed. For instance, the young man on the bus knows that he thinks about offering a seat to the elderly person and what kind of behavior he regards to be more polite. In contrast, he may feel tired because he had a hard day at work without necessarily knowing it. This state of awareness accompanies the Impulsive System, and is described as experiential. It consists of a feeling like being tired, happy, sad, and so forth.

The Reflective and the Impulsive System have a *final common pathway* to behavior represented by motor schemata. Motor schemata are subsumed to the Impulsive System. They comprise frequently co-occurring motor-representations in sensory-motor clusters. Motor schemata are activated by input of the Reflective and the Impulsive System and elicit overt behavior if a given threshold is exceeded. Depending on the compatibility of the motor schemata, the Reflective and the Impulsive System may interact *synergistically or antagonistically*. For instance, participants judged foreign statements as more

convincing, when they nodded rather than shook their head. This was true even if the nodding and shaking was disguised as testing headphones for use on dance floors (Wells & Petty, 1980). In most cultures, nodding is a nonverbal signal for agreement. Therefore, the Impulsive System associates motor-schemata for nodding with agreement behavior. Consequently, nodding strengthens the persuasive power of arguments that are processed within the Reflective System and makes the arguments more convincing. On contrary, shaking one's head is associated with disagreement, and, therefore, weakens the persuasive power of arguments. The ways of interaction between the Impulsive and the Reflective System are manifold, and may take place at every step of information processing. However, there is an asymmetry such that the Reflective System always involves the activation of the Impulsive System, whereas the Impulsive System is able to process information without inferences from the Reflective System.

To summarize the characteristics of both systems, the Reflective System generates knowledge through propositional categorization and allows for the intentional control of behavior. In contrast, the Impulsive System represents an associative network that binds together frequently co-occurring perceptual or behavioral features without any intentional controllability. Nevertheless, reflective operations may have an effect on the Impulsive System. Since information processing in the Reflective System is based on elements that are retrieved from the Impulsive System, reflective operations also influence the associative links in the Impulsive System. As a consequence, frequent propositional categorizations reorganize the associative store and form associative clusters that differ in abstractness. Thus, the associative clusters may represent concrete perceptual concepts or more abstract semantic concepts or schemata. However, the clusters are not assumed to comprise any semantic meaning by themselves, and their elements are only related due to frequently co-occurring activation.

Returning to the conceptualization of explicit and implicit representations, the Reflective-Impulsive Model is convenient to elaborate on their specific characteristics. *Explicit representations* correspond to the propositional categorizations of the Reflective System, that is, explicit representations consist of concepts that are linked by assigning a truth-value to their relationship. Therefore, explicit representations are introspectively accessible. *Implicit representations* correspond to the associative clusters of the Impulsive System, that is, implicit representations consist of concepts that are linked as a result of frequent co-activation. Therefore, implicit representations are accessible only through

procedures that are sensitive for the effects of frequent co-activation. Indirect measures are assumed to represent such procedures.

The manifold interaction between the Reflective and the Impulsive System does not imply that explicit and implicit representations are always different from each other with respect to the content of information they comprise. However, explicit and implicit representations are always different with respect to the form in which information is made available. To illustrate the differences between explicit and implicit representations a painting may help, e.g. Caravaggio's "Amor Victorious". Consider a thought experiment in which a person goes in an art gallery, looks at the painting, and tries to make notes. Writing a description of the painting stands for explicit representations. Preparing a pencil drawing of the painting stands for implicit representations. The pencil drawing, if it's well done, is a fairly analogous representation of the painting. That means, the pencil drawing represents the objects and their locations as they are on the painting, e.g. that Cupid sits on a bed next to a crown, upon a celestial globe, with music instruments and pieces of a body armor next to his feet. In contrast, the written description may list all of these things, and additionally tell that Cupid celebrates a triumph over the symbols of power, science, art, and glory. Thus, the written description is a fairly abstract representation of the painting.

In order to elaborate this metaphor, consider that the person prepared both a pencil drawing and a written description of the painting. At home, the person tells a friend about the painting and shows her or him the drawing and the text. The drawing gives a direct impression about the original whereas the text gives useful comments. The text may be improved from looking at the drawing, but it is rather hard to improve the drawing only with the information provided in the text. The relation between the drawing and the text stands for the interactions between explicit and implicit representations. Explicit representations, the text, consist of concepts that are retrieved from implicit representations, the drawing, and that undergo a process of reasoning and intending. The outcome of this process is a series of propositions, i.e. clear statements about what and why is depicted on the painting. The friend who reads the text and looks at the drawing stands for a psychologist who employs either direct questionnaire measures or indirect assessment procedures. Obviously, the best thing is to use both.

2.3 Direct and Indirect Measures of the Personality Self-Concept

There is confusion about a common terminology for direct and indirect measures (Fazio & Olson, 2003). To resolve the confusion, in this work, the terms explicit and implicit representations are used as labels for the constructs, whereas the procedures to assess these constructs are labeled as *direct* and *indirect measures*, respectively. It should, nevertheless, be noted that, in current literature, direct measures are also referred to as explicit measures, and indirect measures as implicit, unobtrusive, non-reactive, or projective measures. Direct measures openly ask individuals to inform about their thoughts, feelings, and behaviors. In contrast, indirect measures draw inferences from the individuals' reactions in different types of tests and procedures.

Whether direct or indirect, both measures have to meet *psychometric criteria* to serve as instruments that are apt to assess interindividual differences. Psychometric criteria refer to aspects of objectivity, reliability, and validity. Objectivity indicates the independence of a measure from situational effects. Reliability refers to the internal consistency or test-retest stability of a measure. Validity informs about what is assessed or predicted by a measure. (Different aspects of validity are discussed in Chapter 2.5.)

Examples of *direct measures* to assess different aspects of the personality self-concept are manifold, for example, the Revised NEO Personality Inventory from Costa and McCrae (1992). Direct measures are based on verbal self-report and rely on information that is intentionally given to inform about the self. In various domains, direct measures were shown to possess satisfactory psychometric properties (e.g., Pervin & John, 2001).

Examples of *indirect measures* are projective procedures, procedures that are based on linguistic effects, and chronometric procedures. *Projective procedures*, like the Thematic Apperception Test (TAT, Murray, 1943), employ the presentation of ambiguous stimuli. Respondents are assumed to project their own thoughts, feelings, and behaviors onto these stimuli. Projective procedures are criticized to be overly susceptible to contextual influences, and to show poor or moderate inter-rater reliability, as well as low reliability and validity (e.g., Aiken, 1996). Additionally, projective procedures are usually very time consuming.

A procedure that explores *linguistic effects* is the Adult Attachment Interview (George, Kaplan, & Main, 1985). In this interview, the detailed and specific report of experiences with one's own parents indicates secure rather than insecure attachment styles.

The interview is very time consuming. Nevertheless, the results of interviewed parents show good predictive validity for the attachment behavior of their child (van Ijzendoorn, 1995).

Chronometric procedures are based on response latencies. Examples of chronometric procedures are priming methods (e.g., Fazio, Sanbonmatsu, Powell, & Kardes, 1986) and the Implicit Association Tests (Greenwald et al., 1998). Priming methods explore whether the presentation of a stimulus, that is, the prime, influences the speed of response to a different stimulus, that is, the target. Priming methods were shown to be valid for the study of sample means and group differences (for a review, see Fazio & Olson, 2003). However, priming methods reach only low effect sizes, and show small to moderate reliability at best (e.g., Kawakami & Dovidio, 2001). In contrast, the Implicit Association Tests or IATs (Greenwald et al., 1998) were shown to meet psychometric criteria for the assessment of the personality self-concept (e.g., Asendorpf et al., 2002; Egloff & Schmukle, 2002). I refer to IAT measures in plural to make it clear that they represent different applications of a general procedure rather than a specific test (cf. Fiedler, Messner, & Blümke, 2003). The general IAT procedure is described in detail in the following section.

Although indirect measures revealed weaker psychometric qualities than direct measures in most cases, indirect measures were always a matter of enormous interest in psychological research (for a review, see Fazio & Olson, 2003). The reasons for this fascination refer to *two limitations of direct measures* (Greenwald & Farnham, 1995). First, direct measures rely on verbal report that is intentionally given to inform about the self. Therefore, direct measures are susceptible to self-presentational biases. Second, direct measures rely on representations of the personality self-concept that are accessible through introspection. Therefore, direct measures may not reflect the entirety of an individual's knowledge about his or her personality.

Altogether, direct measures of the personality self-concept aim to assess the knowledge about one's personality that is embodied in explicit representations. Indirect measures aim to assess the knowledge about one's personality that is embodied in implicit representations. The next section describes an indirect chronometric procedure, the Implicit Association Tests, in more detail.

2.4 Implicit Association Tests (IATs)

This section deals with the Implicit Association Tests (IATs) that had an enormous impact on psychological research since the initial publication five years ago (Greenwald et al., 1998). The IATs are referred to in plural to indicate that they represent a general measurement procedure rather than a specific test. The first section of this section presents the sequence of tasks that is realized by all IATs. In the second section, different accounts for the effects of IATs are discussed.

2.4.1 The Procedure of IATs

Implicit Association Tests are designed to compare speed of response between two different pairings of a double discrimination task. One discrimination task asks for the categorization of a binary *target concept*, for example, 'flower' versus 'insect' The other discrimination task asks for the categorization of a binary *attribute concept*, for example, 'positive' versus 'negative'. An IAT pairs both categorizations within a double discrimination task, and implements the two possible pairings. One pairing requires one response for one target and one attribute category, and another response for the alternative target and the alternative attribute category. The other pairing leaves responses for the attribute categories the same but exchanges the responses for the target categories.

An IAT starts by introducing participants to the target, and, subsequently, to the attribute concept. For instance, an IAT that assesses attitudes toward flowers and insects *first* trains participants to press the left response key when a flower name is presented on the screen and the right response key when an insect name is presented on the screen (see Table 1). In the *second* sequence, participants are trained to press the left key for positive words and the right key for negative words. The *third* sequence combines the target and the attribute discrimination, and asks participants to respond left to flower names or positive words, and right to insect names or negative words. The *fourth* sequence reverses the target discrimination, and assigns the left response to insect names and the right response to flower names. Finally, the *fifth* sequence combines the attribute and the previously reversed target discrimination, and asks participants to respond left to insect names or positive words, and right to flower names or negative words.

Table 1

Task Sequence and Stimuli of an Implicit Association Test to Measure Attitudes toward Flowers and Insects

			Response key assignment	
Sequence	Task		Left key	Right key
1	Target discrimination		Flower	Insect
2	Attribute discrimination		Positive	Negative
3	Initial combined task		Flower, positive	Insect, negative
4	Reversed target discrimination		Insect	Flower
5	Reversed combined task		Insect, positive	Flower, negative
	Target concept		Attribute concept	
Categories	Flower	Insect	Positive	Negative
Sample stimuli	aster	fly	caress	abuse
	hyacinth	cockroach	freedom	crash
	crocus	mosquito	health	filth
	iris	wasp	love	murder
	rose	termite	peace	sickness

Note. Sample stimuli correspond to Greenwald et al. (1998).

For the calculation of IAT scores, or IAT effects, only response latencies within the combined tasks are relevant. Various variants of IAT scores are based upon the difference in mean response latencies in sequence 5 minus sequence 3. Thus, if participants are quicker in combining flower names + positive words and insect names + negative words relatively to the reverse pairing, they attain low latencies in sequence 3 and high latencies in sequence 5. This would result in a positive IAT score. Normally, participants evaluate flowers more positively than insects on direct attitude measures (Greenwald et al. 1998). This was equally indicated in the indirect measure by a positive IAT effect. Greenwald and colleagues (1998) concluded that quicker responses plausibly reflect *stronger associations* for flower + positive and insect + negative *relatively* to flower + negative and insect + positive. The combined task that reveals quicker responses in most respondents is often referred to as the *'compatible'* task. Thus, in the flower-insect attitude IAT, the flower + positive and insect + negative pairing would represent the 'compatible' task.

The conventional IAT scoring algorithm was presented in the initial publication of IAT data (Greenwald et al., 1998). This procedure discarded training trials from the combined blocks, and was based on log-transformed latencies. Recently, Greenwald, Nosek, and Banaji (2003) proposed an improved algorithm for IAT scores that are referred to as *D measures*. D measures (a) employ untransformed response latencies from all trials of the combined blocks, (b) include a latency penalty for error trials, and (c) are individually calibrated by each respondent's standard deviation of latencies. D measures outperformed the conventional IAT scores with regard to several criteria. In contrast to conventional scores, D measures were more resistant to contamination by response speed differences, and less affected by prior experiences with the IAT procedure. D measures are also yielded in larger effect sizes and higher correlations with direct self-report measures.

One limitation that results from the procedure of IATs is that it is confined to *relative* association strength: An IAT effect reflects the association strength of one pairing of target and attribute categories relatively to the reverse pairing. For instance, a positive flower-insect IAT score in the above example merely reflects that one evaluates flowers more positively, or less negatively, than insects. This does not illuminate whether one endorses either positive or negative attitudes toward either flowers or insects. Thus, IATs assess associations between an attribute concept and a target category only in relation to an opposing target category.

Therefore, alternatives to the IAT were developed to allow for single target categories, that is, the EASTs ("Extrinsic Affective Simon Tasks", De Houwer, 2003a), the EMAs ("Evaluative Movement Assessments", Brendl, Markmann, & Messner, 2003), the GNATs ("Go/No-Go Association Tasks", Nosek & Banaji, 2001), and the STIATs ("Single Target IATs", Wigboldus, 2003). A variant of the EMA, the Indirect Association Procedure (IAP) was developed in Study 1 to assess the implicit self-concept of shyness. This procedure is described in the pilot studies of Study 1. The other procedures are not discussed in more detail because they are not directly related to this research. The common goal of all of these measures is to assess associations between concepts by contrasting opposing pairings of the concepts.

A second limitation of the IAT is that it may not be unquestionably qualified as an *indirect* or an unobtrusive measure. Indirectness usually refers to (a) unawareness, and therefore (b) uncontrollability of what is measured by a certain procedure (Greenwald & Banaji, 1995). However, the first aspect, unawareness, is not true for IATs as they

explicitly introduce the target and the attribute concept. Concerning the second aspect, uncontrollability, empirical evidence shows that IATs can be both robust against (Banse, Seise, & Zerbes, 2001; Egloff & Schmukle, 2002; Kim, 2003) and susceptible to (Fiedler & Blümke, 2003) volitional influences. Nevertheless, IATs were fakable only when participants were informed beforehand how the calculation of the IAT score works (Fiedler & Blümke, 2003). In addition, IAT results can be influenced by mind sets of the participants that they more or less deliberately acquire before the test (see the special issue of the *Journal of Personality and Social Psychology, 71*, 2001). Among the new tests only the EAST (De Houwer, 2003a) does not explicitly introduce the target concept. In this work, IATs are labeled as indirect measures because they aim to assess implicit representations. However, this does not imply that the procedure and the outcome of IATs are necessarily unaware and uncontrollable.

The third limitation of IATs is that they do not allow for the simultaneous assessment of *multiple target or attribute concepts*. Particularly in research on personality differences, one is often interested in simultaneously assessing numerous personality-describing attributes with the IAT, as it is possible in direct questionnaire measures. Among the new tests, the EMA (Brendl et al, 2003) and the EAST (De Houwer, 2003a) allow for multiple concepts although right now empirical evidence is lacking that these procedures assess multiple implicit concepts without major confounds between them.

2.4.2 Accounts for the IAT effect

IATs operate on the basic premise that it is easier to pair two highly associated concepts in one response than to separate them in different responses (Greenwald & Nosek, 2001). However, this does not elucidate (a) *how* the pairing of associated concepts facilitates the response, (b) whether this is uniquely driven by *association strength* or by other aspects of conceptual propinquity, and (c) what the *method-specific influences* of the IAT are. In contrast to the manifold research on the validity of IATs (for reviews see, e.g., Fazio & Olson, 2003; Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2003), only a few studies have looked at the underlying cognitive mechanisms that produce the IAT effect.

Dasgupta and Greenwald (2001) accentuated the similarities between IATs and evaluative priming (Fazio et al., 1986). According to the authors, both kinds of tasks are based on the assumption that attitudes are activated automatically, and, therefore, facilitate

the processing of evaluatively congruent stimuli. For both tasks, the strength of response facilitation is considered to be a measure of the strength with which the attitude object is automatically associated with a positive or negative evaluation. However, Mierke and Klauer (2001) outlined differences between IATs and priming, considering both the semantic and the evaluative priming task. These authors mentioned that the *spreading activation* account that was shown to explain semantic (Neely, 1991) and evaluative (Bargh, Chaiken, Raymond, & Hymes, 1996; for a different explanation, see De Houwer, Hermans, Rothermund, & Wentura, 2002; Klauer & Musch, 2003) priming effects is incapable of explaining IAT effects. Originally, the spreading activation theory was used to describe information processing in semantic networks as a spread of activation between interconnected nodes that represent units of conceptual knowledge (Collins & Loftus, 1975). Given this conceptualization, the spreading activation model is unable to explain differences between the compatible and the incompatible IAT pairing, because both are identical with respect to stimulus composition, and, thus, also with respect to stimulus-triggered activation patterns (Mierke & Klauer, 2001).

Nevertheless, spreading activation may still be an appropriate metaphor for describing differences between IAT pairings in regards to the broader view of the two-systems model of Strack and Deutsch (2003). According to this model, spreading activation is the universal method of information processing within the Impulsive System, and provides, more or less, direct links between perceptual inputs and motor schemata. These links may be offered more easily if associated inputs - that is, inputs producing somehow similar activation patterns - are matched to identical motor schemata. In contrast, if unassociated inputs - that is, inputs producing different activation patterns - have to be matched to identical motor schemata, direct stimulus-response links may be hindered.

This view corresponds to the assumption of learned associations between the response keys and the assigned attribute category (Neumann et al., 1998) as well as to the *stimulus-response compatibility* mechanism (De Houwer, 2001, 2003b). This mechanism argues that there is a compatibility between stimulus and response in the compatible pairing because responses are unambiguously associated with an evaluative or semantic meaning. In contrast, stimulus-response compatibility is missing in the incompatible pairing because response representations are ambiguous with respect to a certain meaning. According to De Houwer, stimulus-response compatibility stems from the *relevant feature* rather than from the *irrelevant feature* of target exemplars. The relevant feature reflects the

assignment of the target exemplars to a target category, for example, “flower” or “insect”. The irrelevant feature reflects the overlap of the target exemplars with an attribute category, for example, “positive” or “negative”. Although the relevant and the irrelevant features are perfectly confounded in typical IATs, the relevant feature of target exemplars seems to be also relevant for the IAT effect. De Houwer (2001) showed that the positive or negative valence of target exemplars had little or no impact on the IAT effect. Therefore, he employed an IAT that assessed the attitudes of British participants towards British versus foreign names. This IAT revealed a preference for combining British names with positive attributes that was not distorted by the valence of British (e.g., Princess Diana or Margaret Thatcher) and foreign (e.g., Albert Einstein or Adolf Hitler) names.

The stimulus-response compatibility model may also account for the frequently replicated finding that the IAT effect is smaller if the incompatible pairing is completed before the compatible pairing (Greenwald, Nosek & Banaji, 2003). If participants are first trained in the incompatible pairing that there is no contingency between the response keys and evaluative or semantic meanings, this has to be extinguished in the compatible pairing. In contrast, if participants first learn that there is stimulus-response compatibility and afterward have to ignore stimulus-response incompatibility, they show larger IAT effects. Recently, the effect of task order was shown to be considerably reduced if additional trials in the reversed target discrimination (sequence 4 in Table 1) were added (Nosek, Greenwald, & Banaji, 2003). Thus, the impact of a preceding compatible or incompatible pairing on the second combined task seems to be minimized, if participants spend more time training the reversed target discrimination. This provides further evidence for a stimulus-response compatibility model, that is, the influence of learned associations between the concept categories and the response keys.

Of course, whether a pairing may be referred to as compatible or incompatible depends on the responses of the participant. Usually, the ‘compatible’ pairing is operationally defined as the IAT task that is completed the quickest for the majority of participants. Therefore, differential effects of task order are not only relevant to interpret IAT scores for participants with different task order, but also for participants with the identical task order. Importantly, for participants with positive IAT scores, that is, participants that are quicker in the first and ‘compatible’ pairing, the differences between the compatible and the incompatible pairing are maximized through the task order effect.

For participants with negative scores, that is, participants that are quicker in the second and ‘incompatible’ pairing, the differences between the pairings are minimized.

Consequently, if one is interested in employing the task order effect to maximize differences within a group of participants, the pairing which is compatible for the respective group should be placed first. For instance, if one explores differences between shy and moderately shy participants, that is, participants which are all quicker in combining ‘me’ + ‘shy’ and ‘others’ + ‘nonshy’ relative to the reverse pairing, the ‘me’ + ‘shy’ pairing should be put first. However, if one explores differences that are symmetrically distributed around zero, the differential effect of task order should be removed through additional trials in the reversed target discrimination (Nosek et al., 2003). Anyhow, if interindividual differences in addition to cross-group groups differences are the matter of interest, the task order should never be counterbalanced across participants. Otherwise, order variance is confounded with interindividual variance.

Karpinski and Hilton (2001) suggested that IATs are influenced by *environmental associations*. These authors employed an IAT to explore the malleability of attitudes toward youth and elderly. When participants were exposed to youth + negative and elderly + positive word pairings, the IAT effect was less biased toward youth + positive. Direct attitude measures, in contrast, were unaffected by the manipulation (Karpinski & Hilton, 2001, Study 3). The authors concluded that IATs reflect associations one has encountered in environment. These associations, however, do not reveal personal attitudes of participants. Importantly, in this study, the IAT effect was only modified by the manipulation but not completely reversed. Moreover, even if IATs are susceptible to learning experiences, this may also indicate the effects of these experiences on the individual’s implicit attitudes rather than merely environmental associations (cf. Dasgupta & Greenwald, 2001).

Mierke and Klauer (2001, in press) reported a *task-switching* account of the IAT effect that also explains method-specific variance in IATs. The model states that attribute-related information is sufficient for fast and accurate responding within the compatible condition. Therefore, in this condition participants neglect to switch between target-based and attribute-based decision on a substantial proportion of trials. As participants neglect to switch, they also avoid task-switching costs. However, task-switching costs cannot be evaded and therefore affect response latencies in the incompatible condition. Consequently, Mierke and Klauer (2001) showed that switching between target and

attribute discrimination produced significantly more costs in the incompatible than in the compatible IAT pairing.

More importantly, task-switching performance was also shown to represent stable interindividual differences in another set of experiments. Mierke and Klauer (in press) demonstrated that IAT effects could be obtained with an IAT that was not based on pre-existing associations between targets and attributes. In that instance, the IAT experimentally imposed a contingency between the target category (color) and the attribute category (size) of geometrical objects, so that all blue objects were big and all red objects were small. The geometrical objects IAT revealed an internally consistent IAT effect that correlated even with the absolute scores of an extraversion IAT, $r = .39$. The correlation was calculated using absolute scores because interindividual differences in task-switching performance were expected to predominantly affect the incompatible IAT pairing. Whether a pairing is incompatible, however, is a function of a participant rather than a function of an IAT. Thus, participants with poor task-switching performance slow down their responses in the incompatible pairing, and add an extremity bias to their IAT scores. This extremity bias is better represented by absolute scores rather than by IAT raw scores. Since no participant showed negative scores in the geometrical objects IAT, absolute scores and raw scores were identical for this IAT. In sum, the correlation between the geometrical objects IAT and the extraversion IAT could not be interpreted in terms of convergent validity, and indicated a reliable contamination of both IATs with method-specific variance.

Interestingly, the correlation between the geometrical objects IAT and the extraversion IAT was rendered not significant when IAT scores were computed as D measures. D measures are individually standardized for latency variability and refer to the improved scoring algorithm from Greenwald et al. (2003) (see Chapter 2.4.1). This individual calibration seems to control for method-specific variance that is produced by task-switching costs.

Brendl, Markman, and Messner (2001) suggested a *random walk model* with variable response thresholds in order to explain IAT effects. The model posits that information on incoming IAT stimuli is accumulated until a certain response threshold is reached. In the compatible pairing, valence and concept information on target stimuli contributes simultaneously to reach the response threshold. In the incompatible pairing, valence information and concept information on target stimuli are conflicting, and

contribute oppositely to reach the threshold. According to the authors, this leads to a *criterion shift* and higher response thresholds for targets and attributes, thus, to generally slower responses within the incompatible condition. However, the notion of a general criterion shift would not explain differential effects of task-switching that were reported by Mierke and Klauer (in press).

Rothermund and Wentura (2001, 2003) suggested a *figure-ground model* of IAT effects. According to this model, the two target categories as well as the two attribute categories differ with respect to salience. The salient category of a target and an attribute concept serve as "figure" on the "ground" of the opposing nonsalient category. During the compatible IAT pairing, both salient categories are mapped to one response key, and both nonsalient categories to the other response key. Therefore, participants can base the discrimination of categories on the figure-ground information alone. In a series of different experiments, Rothermund and Wentura (2001, 2003) dissociated effects of salience from effects of association strength, and showed that salience asymmetries may produce IAT effects. The authors concluded, that IAT effects do not necessarily rely on associations between categories. However, this does not rule out that associations may produce IAT effects as well, and that salience asymmetries themselves may be the result of associations. For instance, in the flower/insect IAT, insects may be the salient category because they are associated with negative valence. Thus, salience asymmetries may simply reflect different associations of flowers and insects with positive and negative attributes. Furthermore, Mierke and Klauer (in press) showed that salience asymmetries, as well as associations, are not a necessary precondition for IAT effects.

Steffens et al. (2003) proposed a *two-factor model* and classified previous accounts of the IAT effect into those that are concept-based and those that are stimulus-based. Concept-based accounts concentrate on target-attribute associations at the concept level, whereas stimulus-based accounts concentrate on individual features of target and attribute exemplars. Steffens et al. (2003) concluded that both accounts contribute to the IAT effect, and labeled the former as task factor, and the latter as stimulus factor. The task factor and the stimulus factor are similar to the relevant and irrelevant feature account from De Houwer (2001, 2003b), respectively, but they refer to features of both, the target and the attribute concept. The *task factor* accounts for a simplified task representation throughout the compatible IAT pairing because of a dimensional overlap (Kornblum, Hasbroucq, & Osman, 1990) between the target and the attribute concept. Thus, in the compatible

pairing, participants do not need to base their discriminations on target and attribute information, but may simply employ the overlapping dimension (e.g., valence). Therefore, participants are faster in the compatible than in the incompatible pairing.

The *stimulus factor* accounts for a modification of the task factor because of consistent or inconsistent cross-category associations. For *consistent cross-category associations*, there is a dimensional overlap between the target exemplars and the attribute concept, or between the attribute exemplars and the target concept, that goes beyond the dimensional overlap between the two concepts. For *inconsistent cross-category associations*, there is a dimensional overlap between the exemplars and the nonrelevant concept, which is the opposite of the dimensional overlap between the two concepts.

To illustrate different cross-category associations, I employ the categories and stimuli that I already discussed about in De Houwer's (2001) experiment. As one may recall, British participants were quicker in pairing the target category 'British name' with the attribute category 'positive' and the target category 'foreign name' with the attribute category 'negative' than in the reverse pairing. However, exemplars of both target categories differed with regards to their associations with the attribute categories. For consistent cross-category associations, these associations equaled the association between the target and the attribute concept, that is, British names represented positive persons (e.g., Princess Diana) and foreign names represented negative persons (e.g., Adolf Hitler). For inconsistent cross-category associations, these associations were in opposition to the association between the target and the attribute concept, that is, British names represented negative persons (e.g., Margaret Thatcher) and foreign names represented positive persons (e.g., Albert Einstein). Thus, cross-category associations of stimuli are described as consistent when they match the compatible pairing of the two concepts, whereas they are described as inconsistent when they match the incompatible pairing of concepts.

The notion of a *dimensional overlap* between targets and attributes corresponds to Fiedler et al.'s (2003) *redundancy model* of the IAT effect. Redundancy arises in a discrimination task if stimuli constantly differ with regard to more than one aspect (Garner, 1969). Due to redundancy in the compatible IAT task, the discrimination of attributes facilitates the discrimination of targets, because the features of both concepts are correlated. However, the dimension in which target and attribute features correlate may be equally described as dimensional overlap. This overlap can occur on both, the concept level (i.e., the task factor) and the stimulus level (i.e., the stimulus factor).

Steffens and colleagues (Steffens & Plewe, 2001; Steffens et al., 2003) conducted several experiments to explore the influence of the task and the stimulus factor. These experiments differed from De Houwer's (2001) experiment with regard to the following points: Steffens and colleagues (a) manipulated the cross-category associations for both target and attribute exemplars, (b) made sure that these manipulations were true at an explicit level for every single participant, (c) explored effects of different cross-category associations both in separate IATs and within mixed IATs, (d) employed IATs that assessed attitudes towards women or Germans, and (e) conducted experiments with larger sample sizes and more trials in the combined IAT tasks. Results showed an influence of the stimulus factor, that is, cross-category associations, in all experiments. As expected, the IAT effect was larger for consistent rather than for inconsistent cross-category associations. However, inconsistent cross-category associations never completely reversed the IAT effects. Therefore, Steffens et al. (2003) concluded that the task factor, that is, the dimensional overlap between the concepts, played a major role in the IAT effect but may be modified by stimulus features. Nevertheless, other authors showed that IAT effects may be even reversed for inconsistent cross-category associations (Blümke & Friese, 2003; Govan & Williams, 2003).

Mitchell, Nosek, and Banaji (2003) and Nosek, Greenwald, and Banaji (2003) reported results that are in agreement with the two-factor model. Mitchell and colleagues (2003, Experiment 1) showed that IATs with identical target exemplars (liked Black athletes versus disliked White politicians) revealed a more positive evaluation for Black athletes or for White politicians dependent on whether participants had to discriminate occupation (athletes versus politicians) or race (Black versus White) in the IAT. These results clearly underline the importance of the task-factor, that is, the concept categories. However, the same authors revealed that the IAT is equally sensitive to individual stimulus features. The effect of a racial attitude IAT was influenced dependent on whether target stimuli were liked Whites and disliked Blacks or disliked Whites and liked Blacks (Mitchell et al., 2003, Experiment 2). This is in line with the results of Nosek and colleagues (2003, Study 3). In this study, the effect of an IAT that assessed attitudes towards homosexuals was less negative when both male-male and female-female couples were used as targets. Therefore, Mitchell et al. (2003) came to the same conclusion as Steffens et al. (2003). The IAT effect depends on both the *category frame*, that is, the

target and attribute categories, and the *individual exemplars*, that is, the features of the individual stimuli.

Concerning practical applications, one should employ stimuli that (a) well represent all relevant aspects of the category frame, and (b) may not be categorized according to concepts that differ from the category frame (Greenwald & Nosek, 2001; Nosek et al., 2003). When these criteria are fulfilled, there is a good reason to assume that an IAT effect relies on what is represented by the category frame. Stimuli features may provide contextual meaning (cf. Nosek et al., 2003). However, stimuli features are unlikely to severely distort the IAT effect.

To summarize the accounts for the IAT effect, the dimensional overlap between targets and attributes seems to play an important role in most of these accounts. The more dimensional overlap exists between targets and attributes, the more similar are the activation patterns that they produce. If similar activation patterns are matched to identical responses in the compatible IAT pairing, responses are facilitated. However, dimensional overlap is just a broader term than association strength, and does not in turn specify the underlying cognitive mechanisms of the IAT. Nevertheless, this notion clarified that IAT effects may not uniquely stem from associations but also from any features that cause dimensional overlap, e.g., salience, similarity, familiarity, and so forth. Salience asymmetries (Rothermund & Wentura, 2003) and stimulus similarity (Mierke & Klauer, in press) were shown to produce IAT effects, whereas familiarity has yet been ruled out as an alternative explanation for the IAT effect (Dasgupta, McGhee, Greenwald, & Banaji, 2000; Ottaway, Hayden, & Oakes, 2001, Rudman, Greenwald, Mellot, & Schwartz, 1999). Concerning method-specific influences, the absolute IAT scores seem to be affected by task-switching costs (Mierke & Klauer, 2001, in press), and the IAT effect depends upon features of both the employed concepts and the individual stimuli (Mitchell et al., 2003; Steffens et al., 2003).

It should be noted that most of these accounts refer to effects on the IAT score. Only a few studies included correlations between an IAT and direct measures or between different IATs. Importantly, correlations between an IAT and direct measures were unaffected or tended to be somewhat higher even if the IAT effect was reduced by procedural variations or stimulus features (Mierke & Klauer, in press; Nosek et al, 2003; Steffens & Plewe, 2001). Task-switching costs did not only affect absolute IAT scores but also reliably contaminated correlations between conventionally calculated absolute IAT

scores (Mierke & Klauer, in press). One should be careful when models concerning the IAT effect are employed to draw conclusions about the correlations of IAT scores (cf. Asendorpf, 1992).

2.5 Multitrait-Multimethod Validation of Indirect Measures

Campbell & Fiske (1959) pointed out the employment of the *multitrait-multimethod matrix* for the validation of personality measures. Multitrait-multimethod validation means that more than one trait as well as more than one method are included in the validation process. Traits and methods are completely crossed in a matrix, such that every trait is assessed with every method. Nevertheless, completely crossed multitrait-multimethod matrices are the exception rather than the rule in the study of interindividual differences (Fiske, 1987). Within completely crossed designs two different aspects of validity, that is, convergent and discriminant validity, are analyzed simultaneously.

Convergent validity is calculated as monotrait-heteromethod correlations, thus, as correlations of a single trait that was measured with different methods. *Discriminant validity* is calculated as heterotrait-monomethod and as heterotrait-heteromethod correlations, thus, as correlations of different traits that were measured with the same and with different methods, respectively. Desirably, the monotrait-heteromethod correlations are larger than the heterotrait-monomethod correlations. If this is not the case, data variance is dominated by method-specific effects rather than by trait-specific effects. Ideally, the heterotrait-heteromethod correlations are zero, indicating an independence of both traits and methods. (I ignore here that traits and methods may be inversely correlated, see Campbell and Fiske, 1956.) Monotrait-heteromethod correlations can than be unequivocally interpreted as convergent validity, and heterotrait-monomethod correlations can be interpreted as method-specific effects.

Consequently, multitrait-multimethod validation implies *contrasting* convergent and discriminant validity. Convergent validity is only accepted when it is higher than discriminant validity, that is, when convergent validity outperforms the variance that methods share while assessing different traits. An example for shared method variance of direct questionnaire measures is their susceptibility to social desirability concerns. An example for shared method variance of indirect chronometric procedures – yet not for the IATs (Greenwald et al., 2003) – are shared interindividual differences in response latency.

Finally, extremity biases may represent a method factor that affects both direct and indirect measures (cf. Mierke & Klauer, in press).

Following the conceptualization of explicit and implicit representations (see Chapter 2.2), direct and indirect measures are assumed to assess constructs that are partially overlapping but not identical. Thus, regarding correlations between direct and indirect measures it is not clear whether these correlations should be interpreted as convergent or as discriminant validity (cf. Greenwald et al., 2003). As a possible solution, direct-indirect correlations may be viewed either as convergent or as discriminant validity depending on whether one aims at assessing the overlap or the disparity between explicit and implicit representations. Nevertheless, to evaluate the specific characteristics of explicit and implicit representations, an effort has to be made to identify variables that are correlated with indirect but not with direct measures and vice versa. Therefore, other methods than direct measures, e.g., behavioral variables, should be included in the validation of indirect measures.

Such an approach was recently chosen by Asendorpf et al. (2002) who showed a *double dissociation* between a direct shyness questionnaire and an indirect shyness IAT for the prediction of shy behavior. The shyness questionnaire uniquely predicted controlled (but not spontaneous) shy behavior, whereas the shyness IAT uniquely predicted spontaneous (but not controlled) shy behavior. Dissociations between direct and indirect measures for the prediction of controlled and spontaneous behavior were also found for racial attitudes (Dovidio, Kawakami, & Gaertner, 2002; McConnell & Liebold, 2001) and for consumer attitudes (Plessner, Wänke, Friese, & Haar, 2003). Thus, the validation of indirect measures should include the study of convergent and discriminant validity using both direct measures and behavioral observations as methods.

Additionally, the validation of indirect measures should comprise more than one trait, to make sure that results are not restricted to a specific trait. More importantly, in order to correctly evaluate the convergent validity of an indirect measure for the assessment of implicit representations, more than one indirect measure is needed. A comparison between different indirect measures is also necessary to judge the method effects of any specific indirect assessment procedure. In sum, not only different traits but also different indirect methods are crucial for the validation of the implicit self-concept.

2.6 Shyness, Anxiousness, and Angriness

The following two sections deal with shyness, anxiousness, and angriness that were used as dependent variables for the validation of the implicit personality self-concept. The constructs were labeled as shyness, anxiousness, and angriness to make it clear that they refer to personality traits, that is, relatively stable response dispositions, and do not describe emotional states in specific situations (cf. Leary, 1991). Chapter 2.6.1 refers to Lazarus' (1991) emotion theory and explains how personal and situational factors interact to develop shy, anxious, and angry behavior. Chapter 2.6.2 describes how shyness, anxiousness, and angriness are related to neuroticism, extraversion, and agreeableness.

2.6.1 Shyness, Anxiousness, and Angriness and Lazarus' Emotion Theory

Lazarus' (1991) emotion theory defines the formation of emotions as a function of a multi-level appraisal process. Appraisals are considered as reflective or impulsive decisions that estimate a given person-environment relationship and evolve a particular emotion. Each emotion is qualified by its unique *core relational theme*. The core relational theme summarizes personal harms or benefits that result from the person-environment relationship. The core relational theme for anxiety is defined as "facing uncertain, existential threat", and for anger as "a demeaning offense against me and mine" (p. 122).

To construct the core relational theme, an appraisal process generates different evaluative patterns that discriminate among emotions. Therefore, the appraisal process involves a set of primary and secondary appraisals. *Primary appraisals* concern the motivational aspects and the personal stakes in a person-environment encounter. Primary appraisals include three components, which are goal relevance, goal congruence or incongruence, and type of ego-involvement. *Goal relevance* refers to whether an encounter affects personal goals. *Goal congruence or incongruence* is concerned with whether the encounter facilitates or thwarts personal goals. *Type of ego-involvement* deals with aspects of ego-identity, e.g., self- and social-esteem, moral values, or life goals.

Secondary appraisals concern options for coping behavior, that is, the prospects to preserve positive or to avoid negative emotional states. Again, secondary appraisals include three components, which are credit or blame, coping potential, and future expectancy. *Credit or blame* derives from knowing who is held responsible for a frustration or a success. *Coping potential* refers to whether an individual can deal with situational demands and whether a situation offers possibilities to actualize personal goals.

Future expectancy refers to the probability of coping behavior changing things for better or worse.

In order to develop an evaluative pattern that specifies an emotion, one doesn't require all six of the appraisal components for each emotion. For instance, all the necessary and sufficient components for anxiety comprise only the three primary appraisal components. For anger, the three primary appraisal components and the blame are essential (see Table 2). Both, anxiety and anger refer to person-environment relationships that are incongruent with their goal. Consequently, anger and anxiety are conceptualized as negative emotions. For *anxiety*, the type of ego-involvement concerns with an existential threat or a threat to self- or social-esteem. According to Lazarus' model, these appraisal components are necessary and sufficient for the construction of anxiety. To experience anxiety, blame is irrelevant, yet the coping options and future expectations are characterized by uncertainty.

Table 2

The Formation of Anxiety or Anger and Resulting Action Tendencies (Lazarus, 1991)

	Anxiety	*Anger*
Primary appraisals		
Goal relevance	(1) If there is goal relevance, *any emotion* is possible.	
Goal incongruence	(2) If there is goal incongruence, *negative emotions* are possible.	
Ego-involvement	(3) If ego involvement concerns threat to self- or social-esteem, …	
	or existential threat, then emotion possibilities narrow to *anxiety*.	then emotion possibilities include *anger* and *anxiety*.
Secondary appraisals		
Blame	(4) Irrelevant	(4) If there is other- or self-blame, *anger* occurs.
Coping potential	(5) Uncertain	(5) If attack is viable, *anger* is facilitated.
Future expectancy	(6) Uncertain	(6) If attack seems successful, *anger* is facilitated.
Action tendencies	Avoidance, escape	Approach, attack

Note. Essential appraisal components are 1 through 3 for anxiety, and 1 through 4 for anger.

For *anger*, the type of ego-involvement concerns with a threat to self- or social-esteem. When others or the self are blamed for the threat to self- or social-esteem, anger occurs. The anger is directed externally or internally, dependent on whether the blame is directed at another person or at oneself. These appraisal components are necessary and sufficient for the construction of anger. However, anger is facilitated, when coping options favor an attack, and if future expectations are positive about the environmental response to attack.

Although the appraisal components are in hierarchical order like in a decision-tree (see Table 2), the model argues that appraisals are not necessarily sequential. Thus, the appraisal process does not represent a step-by-step scan of the reported components in any fixed order. Moreover, the core relational theme of each emotion is identified very rapidly and, possibly, even simultaneously with the appraisal components. Interestingly enough, this construction of emotional meaning may have been evolved through a *reflective*, self-controlled, and abstract cognitive analysis and via an *impulsive*, unconscious process. Thus, in regards to the reflective and impulsive information processing model from Strack and Deutsch (in press; see Chapter 2.2), the appraisal process may be accompanied by both a noetic, that is, conscious, and an experiential, that is, unconscious, state of awareness.

Furthermore, Lazarus' (1991) model embraces behavioral aspects. Behavioral aspects are represented by different *action tendencies* that result from each emotion. The action tendencies that rise from anxiety are avoidance or escape. For anger, the action tendencies are approach or attack (see Table 2). In the model from Strack and Deutsch (in press), these action tendencies are considered to be the aspects of motivational orientation. In both models, action tendencies or motivational orientations and behavior are activated reciprocally. Thus, strong action tendencies are more likely to elicit the compatible behavior. Similarly, any given behavior may reinforce the action tendency and the emotional state consistent with the behavior.

Even though Lazarus' (1991) model deals with the formation of emotional states in different situations, the model does not neglect interindividual differences. Interindividual differences are represented by different *appraisal styles*. An appraisal style summarizes appraisals of person-environment relationships that are consistent across different situations. These transsituationally consistent response dispositions may also be considered as traits. For instance, someone who gets anxious, both, when trying to get around in a new place, and when delivering a speech in front of an audience, may be described as an

anxious person. The feeling of anxiety within both situations refers to an emotional state, whereas the readiness to appraise both situations as a threat to oneself refers to a personality trait, that is, anxiousness.

Thus, Lazarus' emotion theory considers anxiety, and anger as the outcome of a person-environment relationship. Appraisal styles represent the effects of interindividual differences on this person-situation interaction. For instance, individuals high in *anxiousness* consistently tend to appraise person-environment relationships in terms of threat to self- or social esteem or even in terms of existential threat. Similarly, individuals high in *angriness* consistently tend to appraise person-environment relationships in terms of threat to self- or social-esteem, yet, hold others or themselves responsible for their harmful experiences.

Lazarus' emotion theory does not explicitly refer to *shyness*. Nevertheless, Lazarus' theory allows for the inclusion of shyness. According to Asendorpf (1989a), shyness is associated with two types of concern - fear of the unfamiliar and fear of being negatively evaluated by others. Within Lazarus' theory these concerns may be considered as appraisal styles referring to shyness. These concerns may be also part of the appraisal styles for anxiousness (cf. Crozier, 2001). However, the shyness appraisal styles are more situation-specific and refer, in particular, to two kinds of inhibitions, that is, fear of strangers and fear of social evaluation. More importantly, self-descriptions and behavioral observations in shyness-inducing situations provided empirical evidence that shyness was independently elicited by both kinds of inhibitions (Asendorpf, 1989a).

The strength of Lazarus' emotion theory is in its conceptualization of emotions as the result of a person-environment relationship. Thus, the model embraces both individual and situational aspects. In other words, the model refers to emotions as *traits* and as *states*. Moreover, the model includes the adaptational effects of emotions, since it incorporates coping aspects. The weak point of the model is reflected in the appraisal components which do not clearly and fully narrow down the possible emotional outcomes to a particular emotion. For instance, appraisal components for both anxiety and anger may include the experience of threat to self- or social-esteem. Additionally, anger may not only occur due to threat to self- or social-esteem, but also due to threat to bodily integrity. In summary, Lazarus' emotion theory has not yet been empirically tested concerning all of the appraisal components eliciting particular emotions. Therefore, it remains "something of a mystery" (Lazarus, 1991, p. 151) how and why a particular emotion arises. This is

especially the case, as long as the appraisal process itself is not defined as either sequential or simultaneous. Nevertheless, the model stands for the successful application of a psychological stress theory to a person-environment theory of emotions.

2.6.2 Shyness, Anxiousness, and Angriness in Relation to Neuroticism, Extraversion, and Agreeableness

Presently, the most prominent system for the categorization of personality traits is the five-factor, or *Big Five*, model of personality (e.g., John & Srivastava, 2001). The Big Five model proposes that interindividual differences can be classified along five basic personality dimensions that are both broad (i.e., including a maximum spectrum of different traits) and efficient (i.e., managing this with a minimum set of dimensions). These personality dimensions were labeled as neuroticism, extraversion, openness to experience (or intellect), agreeableness, and conscientiousness. Whereas openness to experience and conscientiousness are not directly relevant for the present studies, neuroticism, extraversion and agreeableness are well suitable for the categorization of shyness, anxiousness, and angriness. *Neuroticism* comprises traits like nervous, anxious, and emotional. *Extraversion* describes traits like gregarious, assertive, and outgoing. *Agreeableness* refers to traits like warm, conciliatory, and helpful.

According to the Big Five model, neuroticism, extraversion, and agreeableness are conceptualized as orthogonal, that is, uncorrelated factors. Thus, the factors describe a three-dimensional space that allows for the categorization of shyness, anxiousness, and angriness. *Shyness* represents a combination of neuroticism and introversion, that is, shyness correlates intermediately with both neuroticism and extraversion (cf. Asendorpf, 1989b). However, shyness is independent from agreeableness. *Anxiousness* is highly correlated with neuroticism, moderately correlated with introversion, and uncorrelated with agreeableness. In contrast, *Angriness* is highly negatively correlated with agreeableness and is weakly positively correlated with both neuroticism and extraversion (Ostendorf, 1990).

The correlation pattern of neuroticism, extraversion, and agreeableness with shyness, anxiousness, and angriness also informs about the correlations between shyness, anxiousness, and angriness. Thus, shyness and anxiousness are positively correlated because they are both correlated with neuroticism and introversion. In contrast, shyness and anxiousness do not correlate with angriness, because angriness is weakly correlated with extraversion and strongly correlated with agreeableness. The correlation pattern of

shyness, anxiousness, and angriness facilitates the study of convergent and discriminant validity of direct, indirect, and behavioral measures. In particular, the comparison of shyness or anxiousness with angriness allows to estimate method-specific effects since these traits are expected to be uncorrelated at least at the level of direct questionnaire measures.

The present studies employed shyness, anxiousness, and angriness as traits under investigation in order to explore the similarities and the differences between the explicit and the implicit personality self-concept. In two Pilot Studies, a new indirect measure (the Implicit Association Procedures, IAP) was adapted to assess the implicit personality self-concept of shyness. Study 1 explored the convergent and discriminant validity between the shyness IAP, the shyness IAT, and direct shyness self-ratings for the prediction of shy behavior. Additionally, Study 1 contrasted the fakability of direct and indirect shyness measures. Study 2 investigated the convergent and discriminant validity of an anxiousness IAT and an angriness IAT for direct measures and for the prediction of anxious and angry behavior after emotion inductions. Study 3 explored further method-specific effects of the anxiousness IAT and the angriness IAT that were found in Study 2.

3 Two Pilot Studies for the Adaptation of a New Indirect Measure for Shyness

3.1 Introduction

Depending on the context, a moderate correlation between direct and indirect measures is sometimes interpreted as convergent validity, sometimes as discriminant validity. However, direct measures were considered to assess explicit representations, and indirect measures to assess implicit representations. Explicit and implicit representations were conceptualized as elements of two different, but interacting systems (see Chapter 2.2). Thus, correlations between direct and indirect measures can neither be unambiguously interpreted as convergent nor as discriminant validity. Instead, in order to correctly evaluate the convergent validity of an indirect measure, a correlational analysis with another indirect measure is needed. Also, a comparison between two different indirect measures is necessary to judge the method effects of any specific indirect assessment procedure. Therefore, an additional indirect measure was developed. The measure was adapted to assess the implicit personality self-concept of shyness, and was pre-tested for the purpose of the next study (Study 1).

Priming methods have only partially been shown to be an adequate referent to the IAT from an individual assessment perspective (see Cunningham, Preacher, & Banaji, 2001, for successfully, and Bosson, Swann, & Pennebaker, 2000, for unsuccessfully correlating priming methods with the IAT). As an alternative the Evaluative Movement Assessment (EMA) from Brendl, Markman, and Messner (2003) was adapted to the study of the implicit personality self-concept.

The EMA was designed to employ automatic movement tendencies for the assessment of implicit preferences and motivations. The procedure induces automatic movement tendencies by two joystick movements that represent either approach behavior (pulling the joystick toward a target) or avoidance behavior (pushing the joystick away from a target). In cooperation with Brendl and Messner the EMA was noticeably modified in order to assess the associative strength between the concept of self and attribute concepts (e.g., shy). The modification of the EMA was named the *Implicit Association Procedure (IAP)*. Its main difference to the IAT is that already the response (pulling the joystick toward a target or pushing it away from a target) has its own valence by triggering an automatic movement tendency. Another difference is that it is possible to specify

unipolar target categories (such as self without specifying an opposite category such as others).

The detailed procedure of the IAP is described in the method section. In line with the EMA methodology it was hypothesized that attributes that play an important role in the self-concept could be responded to more quickly with a joystick movement towards oneself than away from oneself. The opposite should be true for attributes that are not associated with the concept of self. The psychometric properties of three different IAP variants for shyness were pre-tested in two pilot studies. The IAP variant that would be considered for further studies was expected to meet the following criteria. *First*, its internal consistency should be at least $\alpha = .70$. *Second*, it should show a substantial correlation with the shyness IAT, that is, at least $r = .40$. *Third*, it should, like the shyness IAT, correlate intermediately with direct self-ratings of shyness, that is, $.30 < r < .50$. *Forth*, it should, like the shyness IAT, not correlate with social desirability. These criteria were explored in the pilot studies.

3.2 Pilot Study 1: The Bipolar and the Unipolar IAP Variant

In Pilot Study 1 a bipolar and a unipolar IAP variant was examined. Their main difference was that the bipolar variant included *Shy* and *Nonshy* words but no Me and Notme words whereas the unipolar variant included *Shy*, *Me*, and *Notme* words but no Nonshy words.

3.2.1 Methods

Participants and design. Participants were 32 (25 female and 7 male) psychology students that received research participation credit for an experiment on computer aided personality assessment. Their mean age was $M = 22.3$ years, with a range from 19 to 29 years. Since the joystick was situated on the right side of the keyboard and was operated with the right hand, we made sure to select only right-handed participants. Due to technical shortcomings of the first joystick that was used, data from 10 participants of the bipolar IAP version and from 7 participants of the unipolar IAP version had to be excluded.

All participants completed (a) self-ratings on bipolar personality-describing items, (b) the bipolar or unipolar shyness IAP, (c) other personality items, (d) the shyness IAT, (e) two social desirability scales, (f) the IAP variant different from (b), and (g) were interviewed about the experiment. The shyness items of the IAPs and the IAT were included as direct ratings in step (a). The application of the unipolar and the bipolar IAP in

step (b) and (f) alternated between participants, such that half of participants completed the bipolar IAP in step (b) and the unipolar IAP in step (f). The other half of participants completed the IAPs in the reverse order.

Direct self-ratings. All direct self-ratings were assessed on the computer and were presented in a fixed random order. In step (a), participants had to rate their *shyness* on 10 bipolar adjective pairs (e.g., "shy 1-2-3-4-5-6-7 nonshy") that were mixed with 30 conscientiousness, intellect, and irritability pairs. Step (c) comprised 28 personality-descriptive items on a 5-point scale (1 = not at all true for me, 5 = completely true for me). Five items referred to *shyness* and were the same used by Asendorpf et al. (2002). In step (e), participants responded to the 39 items of the *social desirability scales* from Lück and Timaeus (1969; English version by Crowne & Marlowe, 1960) and Stöber (1999; without the Item "Have you ever consumed drugs"). These scales contain 16 and 23 items, respectively, and measure socially desirable responding by asking for socially desirable but infrequent or socially undesirable but frequent behaviors on a true-false format. To obtain a score for socially desirable responding items of both scales were aggregated.

Implicit Association Test (IAT). The shyness IAT was identical to Asendorpf et al.'s (2002) studies. Task sequence and stimuli are depicted in Table 3. IAT scores were computed as the difference between mean response latencies in sequence 5 and sequence 3 (see Table 3). These sequences carried out different combinations of the two target categories (*Me* versus *Others*) with the two attribute categories (*Shy* versus *Nonshy*). Thus, high IAT scores represented quicker associations of Me-Shy and Others-Nonshy as opposed to Me-Nonshy and Others-Shy.

Throughout the five discrimination tasks, category labels assigned to the right or left response key were displayed in the right or left upper screen corner, respectively. Response keys were the number "5" of the right-side numeric keypad and the letter "a" on the left side of the keyboard. On each trail, a stimulus word was displayed in the center of the screen. Participants were instructed to categorize the stimulus as quickly and accurately as possible. Responses were recorded using ERTS software (Behringer, 1994). After correct responses the interstimulus interval was 300 ms. After incorrect responses, the stimulus was immediately replaced by the word *FEHLER* (German for error) for 1000 ms, resulting in a 1300 interstimulus interval. Since this study focused on interindividual differences, and I did not want to confound interindividual variance with order variance, the stimulus order was the same for all participants. In the two combined tasks, the stimuli

alternated between target and attribute discrimination. The 10 target and 10 attribute stimuli were randomized in order within 4 blocks of 20 trials. Internal consistency was evaluated across these 4 subtests. Trials with incorrect responses were excluded from analysis, and response latencies above 3000 ms were recoded as 3000 ms. Since the adaptation of the shyness IAP was based on this data reduction procedure (raw instead of log-transformed latencies, inclusion of first two trials of combined blocks), the reported results refer to such procedure.

Table 3

Implicit Association Test for Shyness: Task Sequence and Stimuli

			Response key assignment	
Sequence	N of trials	Task	Left key	Right key
1	40	Target discrimination	Me	Others
2	40	Attribute discrimination	Shy	Nonshy
3	80	Initial combined task	Me, shy	Others, nonshy
4	40	Reversed target discrimination	Others	Me
5	80	Reversed combined task	Others, shy	Me, nonshy

Stimuli			
Me	Others	Shy	Nonshy
I	they	inhibited	uninhibited
self	them	insecure	secure
My	your	daring	daring
Me	you	candid	candid
Own	other	open	open

Note. The original German stimuli can be found in the appendix.

Implicit Association Procedure (IAP). The IAP was based on the Evaluative Movement Assessment (EMA), developed by Brendl, Markman and Messner (2003). Within Pilot Study 1, two earlier EMA versions were adapted to assess the self-concept of shyness. The two shyness IAP variants were similar to the shyness IAT in that they combined discriminations of *Shy* versus *Nonshy* (attribute discrimination) with discriminations of *Me* versus *Notme* (target discrimination). Contrary to the IAT, only *Me* was explicitly shown on the computer screen and no label for alternative targets was given.

Therefore, *Notme* described the nonself-relevant alternatives better than *Others*. However, the main difference to the IAT was that participants responded by moving a joystick instead of pressing an answer key. With the joystick stimuli had to be pushed toward or away from the word *Me* depending on whether the stimuli had to be associated with *Me* or *Notme*. In the two IAP variants of Pilot Study 1 the joystick was situated before the participant, on the right side of the keyboard. The word *Me* was displayed in the center of the screen, whereas stimuli were presented on its right or left side. For stimuli to appear on the right side the joystick had to be pushed to the left, if the stimulus had to be associated with *Me*, and to the right, if the stimulus had to be associated with *Notme*. For stimuli to appear on the left side the opposite was true.

A *bipolar* and an *unipolar* IAP variant were adapted in Pilot Study 1. The task sequence of both is depicted in Table 4. In the *bipolar* version, there was a discrimination of *Shy* and *Nonshy* but not of Me and Notme words. Participants first had to push Shy words toward Me and Nonshy words away from Me. Then, the answer direction was reversed and Shy words had to be pushed away from Me and Nonshy words toward Me. The IAP score was computed as the difference in mean latency between both tasks (sequence 2 minus sequence 1, see Table 4). The Shy and Nonshy words were identical to the IAT and were randomized in order within 10 blocks of 10 trials. Internal consistency was evaluated across 5 subtests with 20 trials each. In the *unipolar* version there were *Me*, *Notme*, and *Shy* but no Nonshy words. First, participants learned to discriminate the target concepts that consisted of three Me (self, my, own) and three Notme (your, them, other) words that were identical to the IAT target stimuli. In the following initial combined tasks, the five Shy words from the bipolar version were added and had to be pushed toward Me. Finally, the answer direction for the Shy words was reversed. The IAP score was computed as the difference in mean latency between both combined tasks (sequence 3 minus sequence 2, see Table 4). Stimuli were randomized in order within 10 blocks of 11 trials. Internal consistency was evaluated across 5 subtests with 22 trials each.

As in the IAT, participants were instructed to respond as quickly and accurately as possible. The correct answer directions for the *Me* words (ME WORDS = TOWARDS ME) and/or the *Shy* words (SHY WORDS = towards ME or SHY WORDS = AWAY FROM ME) were presented in green color in the middle of the upper screen line. During all trials the word *Me* (white letters) with a frame around it was displayed in the center of the screen. Trials began by displaying the stimulus mask XXXX (red letters) for an interval

of 500 ms at the right or left side of the Me. Next, a target or attribute word (red letters) was presented in the same place. The stimulus disappeared when participants moved the joystick clearly in one direction, whereas the reaction time was registered immediately at the beginning of the movement. Reaction time was measured as the time passed from the beginning of the stimulus presentation. After correct responses the interstimulus interval was 600 ms. After incorrect responses the stimulus was immediately replaced by (a) the word *FEHLER* (German for 'error') if the joystick was moved in the wrong direction, (b) the words *ZU LANGSAM* (German for 'too slow') if there was no response after 3000 ms, or (c) the words *ZU FRÜH BEWEGT* (German for 'moved too early') if there was any response during the presentation of the stimulus mask. All error announcements were displayed in yellow in the center of the screen for 200 ms and were followed by the 600 ms interstimulus interval. Within both IAP variants stimulus order was not randomized between participants. All trials with incorrect responses were excluded from analysis. As the presentation of the stimulus stopped after 3000 ms, there were no response latencies longer than that.

Table 4

Implicit Association Procedure for Shyness: Task Sequence of the Bipolar and Unipolar Variant (Pilot Study 1)

			Joystick direction assignment	
Sequence	N of trials	Task	To me	Away from me
		Bipolar Variant		
1	100	Attribute discrimination	Shy	Nonshy
2	100	Reversed attribute discrimination	Nonshy	Shy
		Unipolar Variant		
1	24	Target discrimination	Me	Notme
2	121	Initial combined task	Me, shy	Notme
3	121	Reversed combined task	Me	Notme, shy

Interview. Finally, participants were asked to comment on the experiment and whether they had difficulties with the IAT or the IAPs. In addition, they estimated the *difficulty* of the IAT and the two IAP variants on five-point scales ranging from 1 = easy to 5 = very demanding.

3.2.2 Results and Discussion

Error rates and distribution of test scores. Error rates were for the bipolar IAP $M = 7.9\%$, $SD = 5.2\%$, for the unipolar IAP $M = 4.9\%$, $SD = 4.2\%$, and for the IAT $M = 6.8$, $SD = 4.0\%$. Differences were tested by a 2x3 ANOVA with order (bipolar vs. unipolar IAP at first) as the between-subjects, and test (bipolar IAP, unipolar IAP, IAT) as the within-subjects factor. Results showed no main effect of order, but a marginal main effect of test, and a marginal interaction effect, $F_{(1, 19)} = .72$, *n.s.*, $F_{(2, 38)} = 2.80$, $p < .10$, $F_{(2, 38)} = 3.16$, $p < .10$. Post hoc comparisons with Bonferoni correction ($p < .005$) indicated that when the bipolar IAP was the first test its error rates were higher than for the unipolar IAP, $t_{(11)} = 4.00$, $p < .005$, $d = 1.64$, as well as error rates for the IAT were higher than for the unipolar IAP, $t_{(11)} = 4.05$, $p < .005$, $d = 1.65$. (The effect size d for repeated measures was computed as $\sqrt{2}(M_1 - M_2)/SD$ where SD is the standard deviation of the difference scores; see Cohen, 1988). All other differences were not even marginally significant, all $|t|_{(11)} < 2.30$, *n.s.*. For all three indirect tests, no participant had error rates higher than 19%, and the distributions of the test scores were not even marginally different from a normal distribution, $Z < 1$.

Reliabilities and correlations of indirect and direct measures. As it can be seen in Table 3, the two IAPs only partially met the criteria for a new indirect procedure. *First*, reliability for both IAP variants was satisfactory and comparable to the IAT, although it tended to be lower for the bipolar IAP. Inspection of scatterplots (first test half against second test half) revealed that the somewhat higher reliability of the unipolar version was driven through one outlier. When this participant was discarded from analysis, Cronbach's α decreased to .73 for the unipolar variant, too. However, exclusion of this participant did not affect the correlations of the unipolar IAP. Together, reliability was slightly smaller for the IAP than for the IAT but still on an acceptable level. *Second*, neither of the IAP variants even marginally correlated with the IAT. Although this correlation was somewhat higher for the bipolar IAP, it still did not reach the substantial convergent validity that was expected. Moreover, the two IAPs were only intermediately correlated, indicating small

convergent validity between both variants. *Third*, concerning direct shyness measures, the bipolar IAP showed high correlations, whereas the unipolar IAP tended to correlate only marginally. Thus, the intermediate correlation of the IAT with direct measures was only replicated for the unipolar IAP, while the bipolar IAP showed high convergent validity with direct self-ratings. *Fourth*, like the IAT, the IAPs did not correlate with social desirability. However, this was also true for direct measures, what may very well be a matter of chance finding, as shyness self-ratings are usually correlated with social desirability (Jones, Briggs, & Smith, 1986), and were so in Pilot Study 2. Finally, the two shyness self-ratings were highly correlated, replicating the convergent validity of the bipolar items, which were used in the indirect tests (Asendorpf et al., 2002).

Table 5

Reliabilities and Correlations of Indirect and Direct Shyness Measures in Pilot Study 1

	1	2	3	4	5	6
1. Bipolar shyness IAP (n = 22)	*.71*	.37[a+]	.25	.68***	.58**	-.13
2. Unipolar shyness IAP (n = 25)		*.84*	.13	.33	.38[+]	.31
3. Shyness IAT (n = 32)			*.85*	.37[*]	.27	-.06
4. Bipolar shyness self-rating (n = 32)				*.88*	.75***	-.08
5. Shyness questionnaire (n = 32)					*.80*	.00
6. Social desirability (n = 32)						*.90*

Note. The reliabilities (Cronbach's α) are shown in *italics* along the diagonal. IAP = Implicit Association Procedure, IAT = Implicit Association Test.

[a] n = 21. [+]$p < .10$ *$p < .05$ **$p < .01$ ***$p < .001$.

According to an advice of the EMA authors (C. Messner, personal communication, December, 2000), the unipolar IAP score may be better calculated when considering response latencies for only the *Shy* without the *Me* and *Notme* words. However, this had almost no effect on the results. If reaction times for *Me* and *Notme* words were excluded rather than included, the unipolar IAP's reliability was virtually the same, α = .86 versus α = .84. The correlation with the bipolar IAP – that was completely without *Me* and *Notme* words – was slightly higher, r = .45 versus r = .37. All other correlations tended to be smaller, such as in the correlation with the IAT (r = -.15, versus r = .13), the bipolar

shyness self-rating ($r = .32$, versus $r = .33$), and the shyness questionnaire ($r = .32$, versus $r = .38$). Together, this illustrated, that inclusion of the *Me* and *Notme* trials into the scoring algorithm did not decrease the validity of the unipolar IAP.

Interview. A 2x3 ANOVA with order (bipolar vs. unipolar IAP at first) as the between-subjects and test (bipolar IAP, unipolar IAP, IAT) as the within-subjects factor was performed on the *difficulty estimates* that participants reported for the three indirect tests. Results showed significant main effects for both factors and a marginally significant interaction effect, $F_{(1, 19)} = 4.52$, $p < .05$, $F_{(2, 38)} = 13.74$, $p < .001$, $F_{(2, 38)} = 3.01$, $p < .10$. Post hoc comparisons with Bonferoni correction ($p < .005$) revealed that *across* the two order groups the unipolar IAP was judged as marginally easier when it was the last rather than the first test, $t_{(19)} = -2.93$, $p < .01$, $d = 1.34$. This was not true for the bipolar IAP, $t_{(19)} = -.06$, *n.s.*. The IAT, that was always the second test, was not judged differently between both groups, $t_{(19)} = -1.49$, *n.s.*. Post hoc comparisons *within* the two order groups indicated that when the bipolar IAP was the first test it was judged as more difficult than the IAT and the unipolar IAP, $t_{(11)} = 4.42$, $p < .005$, $d = 1.80$, $t_{(11)} = 3.80$, $p < .005$, $d = 1.55$. In contrast, when the unipolar IAP was the first test it was not judged as more difficult than either the IAT or the bipolar IAP, $t_{(8)} = 2.63$, *n.s.*, $t_{(8)} = .00$, *n.s.*. Neither the bipolar nor the unipolar IAP were judged as more difficult than the IAT when these were the last test, $t_{(8)} = 2.86$, *n.s.*, $t_{(11)} = 1.08$, *n.s.*.

What made the bipolar IAP - at least when it was the first test - more difficult and, as observed before, more susceptible to errors than the unipolar IAP? In the interview, participants reported that they had difficulties to associate the *horizontal* joystick movement to the right or to the left with a movement toward or away from Me. A movement toward versus away from Me could have been more directly associated with a *vertical* joystick movement, that is, with pulling the joystick towards oneself versus pushing it away from oneself. In the unipolar IAP version, the, although horizontal, Me-Notme dimension was continuously practiced by including the Me-Notme words. In both IAP versions, the Me-Notme discrimination might have been additionally difficult because Me-Notme could not be constantly assigned to a movement to the right versus to the left. Thus, the correct movement direction changed depending on whether the stimulus appeared on the right or the left side of the Me. For example, when Shy words had to be associated with Me, the joystick had to be pushed to the *right*, if a Shy word was presented on the left, versus to the *left*, if a Shy word was presented on the right. Whereas the

assignment of response keys stayed constant during the combined tasks of the IAT, the assignment of movement directions in the IAP did not. As a consequence, the IAP required not only a discrimination of categories but also a consideration of presentation side. Both, the horizontal movement to the right versus to the left and its changeable mapping to Me versus Notme might have made the categorization within the bipolar IAP more difficult, especially since this was not trained by the presentation of Me and Notme words.

The task difficulty of the bipolar IAP may also account for its high correlation with direct shyness measures that reached almost the level of the bipolar IAP's internal consistency. Due to the task difficulty, participants might have been forced to react more reflectively rather than spontaneously. Therefore, the bipolar IAP might have been more consistent with the direct measures than with the IAT. Evidence for this assumption was obtained through a 2x3 ANOVA with order (bipolar vs. unipolar IAP at first) as the between-subjects and test (bipolar IAP, unipolar IAP, IAT) as the within-subjects factor that was performed on *mean reaction times* within the tests. Results showed no main effect of order, but a main effect of test, and an interaction effect, $F_{(1, 19)} = 2.11$, *n.s.*, $F_{(2, 38)} = 13.74$, $p < .001$, $F_{(2, 38)} = 6.21$, $p < .01$. Post hoc comparisons with Bonferoni correction ($p < .005$) revealed the same pattern as for the difficulty estimates within the two order groups. When the bipolar IAP was the first test, it was completed more slowly than the IAT and the unipolar IAP, $t_{(11)} = 4.07$, $p < .005$, $d = 1.66$, $t_{(11)} = 6.78$, $p < .001$, $d = 2.77$. All other differences were not even marginally significant, all $|t|_{(11)} < 2.26$, *n.s.*. Thus, when the bipolar IAP was the first test, participants needed more response time than for the other tests that may indicate that their reactions were more influenced by the reflective system. Another reason for the high correlations between the bipolar IAP and direct shyness measures could be that it was not confounded by *task-switching* accounts (Mierke & Klauer, 2001), as there was only a discrimination of Shy-Nonshy but not of Me-Notme. However, one would rather expect shorter instead of longer response latencies in the absence of task-switching (Mierke & Klauer, 2001). Thus, although the reported response latency differences were significant only for the first IAP and the sample size was small in this study, it would be an interesting topic for further research to explore whether correlations between indirect and direct measures increase with task difficulty and reflection time for the indirect test.

Conclusion. The IAPs' satisfactory internal consistency as well as their congruent validity with direct measures showed that the IAPs are an acceptable procedure for the assessment of interindividual differences. Nevertheless, the interview and the correlation pattern made it clear that three main features had to be changed. First, the joystick had to be moved *vertically* rather than horizontally, as this would better represent a Me-Notme dimension. Second, *Shy and Nonshy* words should be included in the IAP, since the bipolar IAP showed higher correlations with the IAT and direct measures. Third, *Me and Notme* words should also be included, because task difficulty seems to be more comparable with the IAT. These changes were realized in Pilot Study 2.

3.3 Pilot Study 2: The Final IAP Variant

In Pilot Study 2 the final IAP variant was examined. It included like the IAT *Shy*, *Nonshy*, *Me*, and *Notme* words.

3.3.1 Methods

Participants and design. Participants were 31 (27 female and 4 male) psychology students that had not participated in Pilot Study 1. They were recruited for an experiment on computer aided personality assessment, and received research participation credit. Their mean age was $M = 21.6$ years, with a range from 19 to 32 years.

All participants completed (a) the shyness IAP, (b) two social desirability scales, (c) the shyness IAT, (d) personality-describing items, (e) a retest of (a), (f) self-ratings on bipolar personality items, and (g) were interviewed about the IAP. The shyness items of the IAP and the IAT were included as direct ratings in step (f). Contrary to Pilot Study 1, there were no direct shyness self-ratings before the indirect tests.

Direct self-ratings and interview. Again, direct self-ratings were assessed on the computer and were presented in a fixed random order. In step (b), participants responded to the *Social Desirability Scales* identical to Pilot Study 1. Step (d) comprised a 32-item self-monitoring scale and a 8-item irritability scale that were not analyzed for the purpose of the present study. Bipolar adjective pairs in step (f) were identical to Pilot Study 1 and included the *shyness* self-rating. The interview at the end of the experiment was the same as in Pilot Study 1.

Implicit Association Test (IAT) and Implicit Association Procedure (IAP). The shyness IAT was identical to Pilot Study1. For the shyness IAP, the main difference to the preceding variants was that the joystick was moved vertically rather than horizontally. The

joystick had to be pulled toward oneself for words that were associated with *Me,* and to be pushed away from oneself for words that were not associated with *Me*. The task sequence for the final IAP version is depicted in Table 6. Identically to the unipolar variant of Pilot Study 1, participants first learned to discriminate the three *Me* and *Notme* words. In the following initial combined task, the five *Shy* and *Nonshy* words from the bipolar variant were added and had to be pulled to or pushed away from the participant, respectively. Finally, the direction for the *Shy* and *Nonshy* words was reversed, assigning *Shy* words to a movement away from the participant and *Nonshy* words to a movement toward the participant. The IAP score was computed as the difference in mean latency between both combined tasks (sequence 3 minus sequence 2, see Table 6). Stimuli were randomized in order within 8 blocks of 16 trials. Internal consistency was evaluated across 4 subtests with 32 trials each.

Table 6

Implicit Association Procedure for Shyness: Final Task Sequence (Pilot Study 2)

			Joystick direction assignment	
Sequence	N of trials	Task	To the participant	Away from the participant
1	24	Target discrimination	Me	Notme
2	128	Initial combined task	Me, shy	Notme, nonshy
3	128	Reversed combined task	Me, nonshy	Notme, shy

Trial presentation was identical to Pilot Study 1, except for the following points. The word *Me* with a frame around – representing the participant – was presented in the center of the lowest screen line (see Appendix). Stimuli appeared above it in the center of the screen. Stimuli and the stimulus mask were displayed in white to make the screen design more comparable to the IAT. The correct answer directions for the Shy (SHY = ME in sequence 2) or Nonshy (NONSHY = ME in sequence 3) words were presented in a subtle red in the left upper corner of the screen and only during the combined tasks. The joystick was located on the table directly in front of the participant, right in front of the

keyboard and the screen (see Appendix). The joystick could be operated with the right or the left hand, allowing for both right-handed and left-handed participants.

3.3.2 Results and Discussion

Error rates and distribution of test scores. Error rates were for the first IAP $M = 5.3\%$, $SD = 4.3\%$, for the retest IAP $M = 4.1\%$, $SD = 3.7\%$, and for the IAT $M = 5.1$, $SD = 3.3\%$. A one-way ANOVA with test (IAP, IAT, retest IAP) as a within-subjects factor revealed that they were not even marginally different, $F_{(2, 60)} = 2.36$, *n.s.*. For all three tests, no participant had error rates higher than 17% and the distributions of the test scores were not even marginally different from a normal distribution, $Z < 1$.

Table 7

Reliabilities and Correlations of Indirect and Direct Shyness Measures in Pilot Study 2

	1	2	3	4	5
1. Shyness IAP	*.82*	.67***	.60***	.39*	.14
2. Shyness IAP retest		*.87*	.47**	.27	.02
3. Shyness IAT			*.83*	.40*	.00
4. Bipolar shyness self-rating				*.93*	-.40*
5. Social desirability					*.83*

Note. IAP = Implicit Association Procedure, IAT = Implicit Association Test.
$N = 31$. $*p < .05$ $**p < .01$ $***p < .001$.

Reliabilities and correlations of indirect and direct measures. The reliabilities and correlations, which are depicted in the first line of Table 7, met the criteria that were expected from the new IAP. *First*, the IAP's internal consistency was completely satisfactory. *Second*, the IAP correlated highly with the IAT. *Third*, it correlated intermediately and as high as the IAT with the direct self-rating. *Fourth*, the IAP did not, similar to the IAT, correlate with social desirability, whereas this was the case for the direct self-rating. Furthermore, the test-retest reliability of the IAP was lower than its internal consistency, which replicated results for the IAT in other studies (cf. Egloff, Schwerdtfeger, & Schmukle, 2003). Finally, the second IAP showed lower correlations with both the IAT and the direct self-rating. A decrease in validity for the second test was

also shown for the IAT (Asendorpf et al., 2002). Together, the correlational pattern of the IAP met all criteria and was highly comparable to the IAT.

Interview. A one-way ANOVA with test (IAP, IAT, retest IAP) as a within-subjects factor revealed that the *difficulty estimates* for the three indirect tests were not even marginally different, $F_{(2, 60)} = 2.39$, *n.s.*. When the same ANOVA was performed on *mean reaction times* a significant main effect emerged, $F_{(2, 60)} = 7.98$, $p < .001$. Post hoc single comparisons with Bonferoni correction ($p < .015$) indicated that the first IAP was completed more slowly than the IAT and the retest IAP, $t_{(30)} = 3.18$, $p < .01$, $d = .85$, $t_{(30)} = 3.64$, $p < .01$, $d = .81$. However, reaction times between the IAT and the retest IAP were not even marginally different, $t_{(30)} = 1.24$, *n.s.*. Since I did not vary the order of the IAT and the IAP between subjects, I could not examine whether the difference between the first IAT and the subsequent IAT was due to learning effects. Nevertheless, when the order of the IAT and the IAP was counterbalanced across participants in the subsequent study (Study 1), their mean response latencies were not even marginally different, $t_{(295)} = 1.59$, *n.s.*. More importantly, the first IAP in Pilot Study 2 was completed significantly quicker than the first bipolar IAP of Pilot Study 1, $t_{(41)} = 4.67$, $p < .001$, $d = 1.46$.

Conclusion. The correlational pattern as well as the difficulty estimates by the participants revealed a correspondence between IAT and IAP. This is also illustrated by the high correlation ($r = .60$) between both tests that reached almost the level of the IAP's retest reliability ($r = .67$). In general, the IAP seemed to be a good candidate for the purpose of replicating results of the IAT and estimating the method-specific variance of both tests.

4 Study 1: Reliability, Validity, and Fakability of a Shyness IAP and a Shyness IAT

4.1 Introduction

Recently, Asendorpf et al. (2002) adapted an IAT to assess the implicit personality self-concept of shyness. They showed that the shyness IAT (a) reliably assesses individual differences that (b) are partly independent from traditional direct self-ratings, and (c) increase significantly the prediction of spontaneous behavior in a realistic social situation. In Study 1, a total of 139 participants were observed in a naturalistic lab situation that induced shyness, and completed an Implicit Association Test (IAT; Greenwald et al., 1998) and direct self-ratings of shyness. The IAT correlated moderately with the direct self-ratings, and uniquely predicted spontaneous (but not controlled) shy behavior, whereas the direct ratings uniquely predicted controlled (but not spontaneous) shy behavior (double dissociation).

The robustness of the IAT against faking was investigated in Asendorpf et al.'s Study 2 through the experimental variation of participants' self-presentation of being non-shy. A control group of 18 females participated in a shyness-inducing role play allegedly to study social perception. Their shyness IAT scores, direct self-ratings of shyness, observer-judged shyness, and coded behaviors were contrasted with an experimental group of 23 females who completed the same procedures except that they were presented as part of a simulated job application procedure and that the participants were instructed to act non-shy in order to "get the job". As expected, the direct self-ratings and the controlled shy behaviors were much lower in the experimental group whereas the shyness IAT scores and the spontaneous shy behaviors were not lower.

The present study was an attempt to replicate the results of Asendorpf et al.'s (2002) Study 2 with a much larger, sex-balanced sample and to extend this approach into four different directions. *First*, the study attempted to replicate the findings for the shyness IAT with a different, new indirect procedure. *Second*, the study explored dissociations between direct and indirect measures of shyness under faking instructions not only with regard to the group means but also with regard to the correlates of these measures. *Third*, the effects of faking were explored also on observer judgments of shyness. *Fourth*, the state dependence of the indirect measures was examined by contrasting both their mean levels and their correlates between participants who completed them either before or after

the shyness-inducing role play. The next sections discuss each research question in more detail.

4.1.1 Research Question 1: A New Indirect Assessment Procedure

A new measurement tool, the Indirect Association Procedures (IAPs), was employed in the present study in order to estimate the convergent validity of the IAT and a different indirect measure for the assessment of the implicit personality self-concept of shyness. The shyness IAP was pre-tested in two pilot studies (see Chapter 3). The final IAP variant showed good internal consistencies, correlated highly with the IAT, and, similarly to the IAT, intermediately with direct shyness self-ratings. The main difference to the IAT is that the IAP induces automatic movement tendencies and already the response has its own valence by triggering approach (pulling the joystick toward oneself) or avoidance (pushing the joystick away from oneself) behavior. The detailed procedure of the IAP is described in Chapter 3.

4.1.2 Research Question 2: Dissociations of Indirect and Direct Measures Under Faking

Job applicants produce more socially desirable self-descriptions than research participants under most conditions (see, e.g., Ones & Viswesvaran, 1998; Rosse, Stecher, Miller, & Levin, 1998). Similarly, laboratory experiments have shown that revealing one's self-descriptions to the public and faking good instructions increase the social desirability of participants' self-descriptions (Paulhus, 1984). These situational effects on the mean social desirability of self-descriptions are commonly interpreted as a threat to the validity of these descriptions. Less often it has been noted, however, that such mean effects do not necessarily imply a lower validity of the interindividual differences in the self-descriptions. If all individuals fake good to the same extent, the rank order of the individuals and hence the validity of the self-descriptions is perfectly preserved. Only if different individuals fake to a different degree (differential faking), the validity is threatened. There is good evidence for substantial differential faking both in job application and in research settings (Ones & Viswesvaran, 1998; Paulhus, 1984; Rosse et al., 1998).

The present study investigated both the main effect of faking good and the effect of differential faking on indirect and direct measures. Faking was studied by contrasting these measures between an experimental group that was instructed to appear non-shy, and a control group that was instructed to act naturally. The between-group difference in the

means informs us about the general faking susceptibility of the indirect versus direct measures. In contrast, the between-group differences in particular correlates of the indirect versus direct measures can be informative about the amount of differential faking.

According to the findings by Asendorpf et al. (2002), Study 1, a moderate correlation close to .40 is expected between the indirect and direct measures of shyness in the control group. To the extent that differential faking occurs, and affects only the direct self-ratings, the direct–indirect correlation should become much smaller in the experimental group. Furthermore, direct shyness is expected to correlate in the control group somewhat negatively with social desirability tendencies because shyness is a somewhat undesirable personality trait (e.g., Jones, Briggs, & Smith, 1986). To the extent that differential faking occurs, this negative correlation should become much stronger in the experimental group because the more participants fake good, the higher will be their social desirability score, and the lower their shyness score. Such a between-group difference is not expected for the correlations between the indirect measures and social desirability tendencies. These correlations should be low in both groups.

Finally, it was expected that the double dissociation between indirect and direct measures with regard to spontaneous versus controlled behavior reported by Asendorpf et al. (2002), Study 1, would be found not only in the control group but also in the experimental group because the direct self-ratings would be less predictive of spontaneous shy behavior and the indirect measures would be less predictive of controlled behavior.

4.1.3 Research Question 3: Validity of Observer Judgments

In Asendorpf et al.'s (2002) Study 1, the observer judgments of shyness correlated .58 with the controlled shy behavior and .48 with the direct self-ratings, but only .35 with the spontaneous shy behavior and .31 with the IAT. Thus, they seem to reflect more strongly controlled behavior. However, the participants in this study were not particularly motivated to control expressions of shyness. Participants who were instructed to fake non-shyness in Study 2 received only slightly lower shyness ratings by observers of their social interaction despite the fact that they talked much more. It is not clear from this pattern of correlational and mean effects what one should expect for differential faking.

It could be that the observers are strongly influenced by participants' self-presentation in the role play as being non-shy; in this case, the strong correlation with the direct self-ratings would be preserved, and the lower correlation with the indirect measures

would decrease even more because it is less susceptible to faking. Such a pattern would suggest that the validity of the observer judgments for participants' true shyness is undermined by the participants' self-presentation in the role play. However, because behavior in role play situations can be faked less easily than answers in a questionnaire, it seems more likely that the participants' true shyness in the role play perspires to the observers to a great extent. In this case, the direct shyness - observer correlation should decrease, and the indirect shyness – observer correlation should be less affected. Thus, the difference between the faking-induced decreases in the self - observer correlations for direct versus indirect measures informs us about the validity of the observer measures.

4.1.4 Research Question 4: State Influences on the Indirect Measures

Research in Spielberger's state-trait anxiety tradition suggests stability of trait anxiety and increase of state anxiety when assessed immediately after anxiety induction (Spielberger, Gorsuch, & Lushene, 1970; Spielberger, Auerbach, Wadsworth, Dunn, & Taulbee, 1973). In line with these results, investigations with the German version of the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) revealed that state affect is a better predictor for affect report regarding closer and shorter periods whereas trait affect is a better predictor for affect report regarding more prolonged periods (Krohne, Egloff, Kohlmann, & Tausch, 1996).

Recently, a study by Schmukle and Egloff (2003) provided evidence that an anxiety IAT was, in contrast to direct state anxiety measures, not influenced by an anxiety induction. Thus, whereas situational or contextual effects on implicit prejudice and stereotypes were demonstrated in several recent studies (Blair, Ma, & Lenton, 2001; Dasgupta & Greenwald, 2001, Lowery, Hardin, & Sinclair, 2001; Rudman, Ashmore, & Gary, 2001; Wittenbrink, Judd, & Park, 2001), state influences on implicit personality self-concept measures have not yet been shown. In order to make sure that the IAT and the new IAP procedure reflect interindividual differences in the enduring self-concept rather than in fluctuating affective states, it is important to show empirically that state influences are negligible.

This is particularly important because earlier studies have consistently found that the retest or parallel test reliability of IATs is lower than the internal consistency of the IAT (Asendorpf et al., 2002; Bosson et al., 2000; Dasgupta & Greenwald, 2001; Egloff, & Schmukle, 2002; Greenwald & Farnham, 2000). This lower retest reliability could be due

to differential learning effects that occur between test and retest (e.g., some participants develop a more efficient cognitive strategy for the more difficult part of the IAT where they must associate self with incompatible attributes whereas others do not develop such a strategy). Alternatively, the lower retest reliability could be due to effects of state changes between the two tests (e.g., when a first shyness IAT is assessed immediately after a shyness-inducing situation and the retest 20 minutes later, the retest correlation could be lowered by the fact that the first IAT was influenced by the actual shyness experienced in the immediately preceding situation whereas the retest reflected more one's enduring self-concept of shyness). This latter interpretation could be ruled out if it could be shown that the IAT is unaffected by state changes. The robustness of both indirect measures was studied with regard to their mean level and their correlates by contrasting them between participants who completed the indirect measure before or after the shyness-inducing role play.

4.2 Design of the Present Study

In order to answer these 4 research questions, the design of Asendorpf et al.'s (2002) Study 2 was extended in two main respects. First, the new IAP was included in addition to the IAT. Second, both females and males were included and sample size was much larger to be able to detect significant differences between correlations. Statistical power considerations suggest that in order to detect significant between-group differences of approximately .30 with one-tailed tests and a power of .80 (Cohen, 1988), a size of N=120 for each group is required. Because I wanted to experimentally vary both participants' faking tendency and the position of the two indirect tests (before/after the role play situation), a complete between-subjects design would include 2 (faking) × 2 (position) × 2 (indirect procedure) × 120 = 960 participants.

To avoid such an unrealistically large study, I (a) restricted the analysis of the position effect to the faking condition which thus required 240 participants, (b) chose only 60 participants for the control group which still provided sufficiently reliable correlations within this condition and a sufficient power for the faking effects, and (c) had each participant complete one indirect procedure before and the other indirect procedure after the role play, with a between-participant variation of the order of the tests, because I assumed that there would be only minimal transfer effects between different procedures. In this way the total sample size was reduced to 2 (position) × 120 + 60 = 300.

Additionally, I included two social desirability scales to study the effects of faking on the responses to these two scales, and interviewed the participants in the faking condition about possible faking strategies in the indirect procedures.

4.3 Hypotheses

Study 1 tested the following hypotheses.

Hypothesis 1 (Main faking effects). Under faking, the social desirability scores increase, and the direct self-ratings of shyness and the controlled shy behaviors decrease. In contrast, the spontaneous shy behaviors and the two indirect measures are unaffected by the faking instruction, replicating Asendorpf et al. (2002), Study 2.

Hypothesis 2 (Main position effects). Whether the indirect tests are completed before or immediately after the shyness-inducing role play has no effects on their mean level.

Hypothesis 3 (Differential coherence). The indirect and explicit self-concept measures are less strongly correlated in the faking condition than in the control condition.

Hypothesis 4 (Differential relation to social desirability). The direct self-ratings correlate more negatively with the social desirability scores in the faking condition than in the control condition. In contrast, both indirect measures do not correlate with social desirability scores under both experimental conditions.

Hypothesis 5 (Robustness of observer judgments to differential faking). Under faking, the correlation of the observer judgments of shyness with the direct self-ratings decreases more strongly than their correlations with the two indirect procedures.

Hypothesis 6 (Double Dissociation). Both indirect procedures uniquely predict spontaneous (but not controlled) shy behavior whereas the direct self-ratings uniquely predict controlled (but not spontaneous) shy behavior when the alternative predictor is statistically controlled, replicating Asendorpf et al., 2002, Study 1.

4.4 Methods

4.4.1 Participants

Participants were 300 university students (150 female, 150 male; age $M = 24.5$ years, range 20-34 years), none of whom were psychology students or had participated in the lab's earlier studies. All participants were claimed to be native speakers of German.

Most participants were approached on the campus of Humboldt University, Berlin. The remaining were recruited using postings at the university buildings.

Following Asendorpf et al.'s (2002) Study 2 procedure, participants were asked to participate in either "a job application procedure" (faking condition, $n = 240$, 120 of either sex) or "a study on social perception" (control condition, $n = 60$, 30 of either sex). In the first case, they were motivated for participation by informing them that the study included a simulated job assessment center and video feedback on their performance. In addition, they were offered DM 20 (approximately US $ 10) for their cooperation in the 1.5 hour study. In the second case, they were motivated by informing them that they would receive individual feedback on their results after the study. In addition, they were offered DM 15 (approximately US $ 7.5) for their cooperation in the 1 hour study.

4.4.2 Assessments and Measures

Overall procedure and design. The overall procedure and design of Study 1 is shown in Table 8. All participants (a) completed an indirect shyness test (either IAT or IAP), (b) judged themselves on bipolar personality-describing items, (c) were video-taped in a shyness inducing role play, (d) completed a different indirect shyness procedure (IAP or IAT), (e) judged themselves on other sets of personality items, (f) completed a retest of (d), and (h) were interviewed about the indirect tests. Participants in the assessment center condition additionally (g) judged themselves on the personality items of step (e) under a honesty instruction and (i) received video feedback on their performance in the role play by the role play partner. The shyness items were identical for both indirect procedures and were included as direct self-ratings in steps (b), (e), and (g). Thus, the first indirect test was completed before the direct ratings. This excluded possible transfer effects from the direct to the indirect measures. The direct shyness ratings, the IAT, the instructions for the faking and control group, and the role play were identical with the procedures in Asendorpf et al.'s (2002) Study 2.

As can be seen in Table 8, there were two between-subject variations: *faking instruction* and *position* of the two indirect tests. Consistent with their invitation, participants received either the faking instruction (assessment center group) or the honesty instruction (social perception group). Invitations were scheduled such that approximately every fifth participant was in the social perception group. Within each group, half of the participants completed first the IAT and later IAP and IAP retest; the other half completed

first the IAP and later IAT and IAT retest. Assignment to the 2 orders alternated between successive participants. Finally, the participants were thanked, asked for permission of analyzing the videotapes (all gave permission), and were promised individual feedback about their results (only participants in the social perception condition). Four months after the study was finished, all participants received a letter explaining the procedures and general findings of the study, and the control participants were invited for a feedback session where they were informed about their individual results.

Table 8

Overall Procedure and Design of Study 1

	Cover story				Duration (Min.)
	Assessment center: Faking instruction		Social perception: Honesty instruction		
(a) Indirect shyness test	IAT	IAP	IAT	IAP	10
(b) Bipolar self-ratings	Shyness, irritability, conscientiousness, intellect				5
(c) Behavior observation	Shyness inducing role play				5
(d) Different indirect shyness test	IAP	IAT	IAP	IAT	10
(e) Direct self-ratings	- Self-monitoring scale - Bipolar items for shyness and irritability - Social desirability scales				12
(f) Retest of (d)	IAP	IAT	IAP	IAT	10
(g) Retest of (e)	Honesty instruction		-		12
(h) Interview about indirect tests	- Problems with IAT or IAP - Answer or faking strategies				7
(i) Video feedback	Role play performance		-		20
n	120	120	30	30	~1.5/1h

Note. IAT = Implicit Association Test, IAP = Implicit Association Procedure.

Instructions. All instructions were identical to Asendorpf et al.'s Study 2 (2002) Upon arrival at the lab, the participants in the *assessment center condition* received the following instruction: "The following assessment center assesses your ability to present yourself as successfully as possible for a position in a company that you are very interested in. An important part of your future job is to present the company as successfully as

possible in interactions with new clients. Therefore, you must be able to warm-up strangers quickly and to avoid insecure behavior because such insecurity could easily make an unprofessional impression." Then, the experimenter explained the different steps of the assessment center and stressed the point that in order to get the job the participant should make a favorable impression in all parts of the assessment, including both the role play and the personality tests. The instruction “Please do not forget to present yourself in a way that you get the job.” was repeated before each set of direct ratings and each indirect test.

The participants in the *social perception condition* were informed that they would participate in a study on social perception, that is, "how you perceive yourself and how others perceive you". After explaining the different steps of the experiment, the experimenter continued "Please describe yourself in all personality tests as honestly and realistically as possible and act in the role play simply as you would do in real life". The instruction “Please do not forget to present yourself as honestly and realistically as possible.” was repeated before each set of direct ratings and each indirect test.

Role play instructions. Before participants of the assessment center condition were shown into the observation room, they were reminded that "it is very important for getting the job that you show in the role play that you can easily and openly approach strangers". In the *control condition*, the participants were informed that "the role play is informative about particular personality characteristics" and that they would be evaluated by their role play partner after the role play. All participants were informed that the role play would be recorded by two cameras. Then, the role play situation was described: "You are an employee in a company. In your company, the boss will be replaced by a new one. This new boss, your future boss, was supposed to meet the present boss now, but unfortunately the present boss is still in another meeting for about 10 minutes. You have been asked to fill in for these 10 minutes and to make the situation as comfortable for your future boss as possible." In the assessment center condition, this instruction was continued: "You should present yourself as favorably as possible. Have in mind that your role play partner will be your future boss." In the control condition, the instruction was continued differently: "Act in the role play just as you would do in real life."

Role play. The role play situation was identical for all participants. In the observation room, an older-looking, unfamiliar, opposite-sex, advanced psychology student, dressed in a business suit, was sitting at a low table. The participant was asked to take place on a chair, that was put at a 90° angle to the confederate's chair. The confederate

was blind to the experimental condition. S/he was trained to play the role of the future boss described in the instruction. The confederate was instructed to act slightly indignant at the delay of the meeting with the present boss and to slightly patronize the participant. This procedure was designed to induce shyness by (a) the unfamiliarity and (b) the status difference of the boss, (c) the assumed evaluation by the boss, (d) the opposite sex of the boss, and (e) the videotaping.

The role play was videotaped with two cameras that were operated from another room using S-VHS recorders. One camera filmed the participant and the confederate from a 45° angle. These tapes were used for behavioral analyses. A different camera directly looked toward the participant and recorded a zoomed-up view of the participants face. When participants interrupted the role play (e.g., by talking about the role play or walking around), the confederate tried to get them back in the role play as quick as possible. The time period until the role play was continued was defined as missing. For the judgments and codings of shy behavior secondary tapes were prepared that contained the first three minutes of noninterrupted role play of each participant.

Implicit Association Test (IAT) and Implicit Association Procedure (IAP). The same procedures as in Pilot Study 2 were used. Following Greenwald et al. (1998) three aspects were modified concerning data reduction. First, latencies below 300 ms were recoded as 300 ms, as well as IAT latencies above 3000 ms were recoded as 3000 ms. Given that in the IAP the presentation of the stimulus stopped after 3000 ms, there were no response latencies longer than that. Second, the first two responses in the combined tasks were not analyzed. Third, calculations of the internal consistencies and the test scores were based on log-transformed latencies to correct for the skewed latency distribution. However, for presentation purposes, descriptive statistics of the IAT and the IAP are reported in milliseconds.

This data reduction procedure was identical to that used by Asendorpf et al. (2002). To maximize comparability between both studies I do not report results for the improved scoring algorithm that Greenwald, Nosek, and Banaji (2003) suggested recently. However, I analyzed both the present data and the Asendorpf et al. (2002) data with this new procedure but found only minimal changes (differences in correlations below .02). The main reason for the minimal between-procedure difference seems to be that the Asendorpf et al. (2002) procedure already included a major feature of the Greenwald et al. (2003) procedure, namely inclusion of the practice trials for the combined tasks into the analyses.

The gain in internal consistency and validity due to this variation from the original procedure used by Greenwald et al. (1998) was larger than the gain due to the remaining features of the Greenwald et al. (2003) procedure.

Direct self-ratings. Concerning bipolar self-ratings in step (b), the same 10 *shyness* items as in Pilot Study 1 were used. These items were mixed with 30 *conscientiousness, intellect, and irritability* items in a fixed random order. In order to minimize transfer effects from the preceding indirect test, the shyness items occurred only among the last 20 items. Self-ratings in step (e) started with a 32-item self-monitoring scale that should again minimize transfer effects from the preceding indirect test and was not analyzed for the purpose of the present study. The scale was followed by the 10 shyness and irritability items of step (b) and concluded with the social desirability scales of Pilot Study 1. The reliability of the direct self-ratings was separately calculated for the assessment center and the social perception condition and was above $\alpha = .84$ in each case.

Interview about the indirect procedures. All participants were interviewed by the experimenter about (a) problems with the IAT or IAP, and (b) whether they used particular strategies during the IAT or IAP in order to decrease error rate, increase speed, or make a favorable impression.

Judgments of shy behavior. Four student judges who were blind to the experimental condition independently rated their overall impression of the participants' shyness. Each minute of the 3-minute secondary tapes was separately rated on a 7-point scale ranging from 7 = "shy" to 1 = "not shy". Beforehand, the judgments were anchored by two examples of extremely shy and extremely nonshy participants from Asendorpf et al.'s (2002) Study 1. For each participant the 12 ratings were averaged The reliability (interjudge agreement) was above $\alpha = .92$ for both conditions.

Codings of shy behavior. Codings were done on a PC using the Computer Aided Observation System (CAOS) software. This program synchronizes video player and PC and registers onset and offset of behavioral codings when the appropriate key is pressed. Codings were carried out for speech duration, body movements, and tenseness of body posture. Following Ekman and Friesen's (1972) classification body movements were coded as illustrators (movements illustrating speech), facial adaptors (self-stimulations of the face), and body adaptors (self-stimulations of the body). For data analysis body movements and speech duration were considered in terms of their relative duration of the 3 minute observation time. For statistical analyses body movement codings were log-transformed to

correct for the skewed distribution. Tenseness of body posture was defined as deviation from a normally relaxed body posture and was coded on a 3-point scale as normal, slight, or strong tension. Using the weights of 0, 1, and 2, the durations of the three tension categories (in % of observation time) were summed, resulting in scores ranging from 0% to 100%. From these 5 variables, indices of spontaneous and controlled shy behavior were computed as in Asendorpf et al.'s (2002) studies by aggregating the *z*-transformed scores of the three spontaneous behaviors facial adaptor duration, body adaptor duration, and tense body posture, and separately the 2 controlled non-shy behaviors speech duration and illustrator duration. Coding reliability was checked by independent codings of 45 participants by another coder; the reliability was satisfactory for all 5 main behavioral indicators, $r > .86$ in each case.

4.5 Results

The first two sections of the Results section report the main effects of instruction (faking versus social perception) and position (before versus after the role play). Then, the effects of these experimental variations on the correlations between direct, indirect, and behavioral measures are explored.

4.5.1 Effects of Instruction and Position on Indirect and Direct Measures

In this section, the main indirect and direct measures are described, and effects of instruction (main faking effects, Hypothesis 1) and position (before versus after the role play, Hypothesis 2) are analyzed.

IATs. For both IATs, the error rates in the two combined tasks were similar to those in Asendorpf et al. (2002), for the first IAT, $M = 5.1\%$, $SD = 3.6\%$; for the second IAT, $M = 4.9\%$, $SD = 3.8\%$. Inspection of the error distributions indicated three extreme scorers (in the faking condition, 1 participant in the first IAT, 25% error, and 1 in the second IAT, 26% error; in the control condition, 1 in the first IAT, 26% error). All other error rates were below 20%. Therefore, the IAT data of these 3 participants were excluded from all analyses. The distributions of the log-based IAT and IAT retest scores were not even marginally different from a normal distribution, $Z < 1$. Their overall internal consistency α, calculated across IAT scores that were separately determined for the trials 3-20, 21-40, 41-60, and 61-80, was .78 for test and .76 for retest and highly similar for all conditions; in particular, it was not lower in the faking condition (.78 in the faking versus .73 in the control condition for test, and .78 versus .63 for retest, respectively). Thus,

internal consistency was acceptable for all conditions although it was slightly lower than in Asendorpf et al.'s (2002) studies. The retest reliability of the IAT was $r = .68$ and thus highly similar to the parallel test reliability of .66 reported by Asendorpf et al. (2002).

Table 9

Summary Statistics and Instruction Effect for the Main Variables

	Faking $n = 240$[a]		Control $n = 60$[b]		Instruction effect $df = 298$[c]		
Variable (range of scores)	*M*	*SD*	*M*	*SD*	*t*	*p*	*d*
IAT	-115 ms	194 ms	-76 ms	169 ms	1.99	.05	.23
IAP	-85 ms	134 ms	-62 ms	142 ms	1.27	.21	.15
Bipolar shyness self-rating (1-7)	1.85	0.59	3.58	1.01	17.3	.001	2.00
- before role play	1.90	0.64	3.62	1.01	16.3	.001	1.89
- after role play	1.79	0.59	3.54	1.03	17.3	.001	2.00
Social desirability score (0-1)	0.85	0.14	0.48	0.17	17.8	.001	2.06
Observer shyness judgment (1-7)	3.72	1.19	4.11	1.26	2.29	.02	.27
Speech duration (%)	85.9	26.3	68.9	24.7	4.52	.001	.52
Illustrator duration (%)	6.22	5.85	4.82	5.97	1.70	.10	.20
Facial adaptor duration (%)	3.39	10.4	5.08	11.8	1.22	.22	.14
Body adaptor duration (%)	35.1	39.7	28. 6	39.9	1.25	.21	-.14
Tense body position (%)[d]	66.9	29.6	54.1	27.2	3.03	.01	-.35

Note. *M* and *SD* refer to raw scores, statistical tests to log-transformed scores in the case of the IAT and IAP latencies and the body movement codings. The effect sizes *d* were defined such that positive scores indicate less shyness in the faking condition.

[a] $n = 239$ for IAT and IAP; [b] $n = 59$ for IAT and IAP.

[c] $df = 294$ for IAT and IAP, $t = \sqrt{F}$ in case of ANOVAs.

[d] Weighted duration of normal, slight, and strong tension.

Effects of instruction, position, and their interaction on the IAT means were tested by a 2×2 ANOVA. A significant effect was found only for instruction, $F_{(1,294)} = 3.97$, $p < .05$. Table 9 indicates that participants had lower IAT scores in the faking condition than in the control condition. Although the effect size was small, it suggested that some

participants might have manipulated the IAT in order to present themselves as nonshy. Therefore, the participants' reports about faking the IAT in the interview with the experimenter at the end of the study were related to their IAT scores. Of the 239 participants in the faking condition, 58 reported attempts of influencing the direction of the IAT outcome. In 57 cases, they reported to bias their results by having taken the perspective of a nonshy job applicant; one other participant reported to have deliberately committed errors. A *t* test contrasting them with the other 181 participants in the faking condition confirmed the hypothesis that they had lower IAT scores, $t_{(237)} = 1.78$, $p < .05$, one-tailed, $d = .23$. When these 58 participants were excluded from analysis, the remaining participants had only marginally lower IAT scores than those in the control condition, $t_{(238)} = 1.44$, $p < .08$, one-tailed, $d = .19$. In terms of untransformed reaction times, the mean IAT score was –154 ms for fakers, -103 ms for assumed nonfakers, and –76 ms for control participants. Because some of the assumed nonfakers might have tried as hard as the fakers to influence the IAT, but did not report it, the instruction effect for the IAT seems to be due to the tendency of a minority of the participants to take the perspective of a nonshy job applicant.

IAPs. For both IAPs, the error rates in the combined tasks were similar to those in the IAT (for the first IAP, $M = 5.0\%$, $SD = 5.3\%$; for the second IAP, $M = 3.8\%$, $SD = 3.5\%$). Inspection of the error distributions indicated two clear outliers (in the faking condition, 1 participant in the first IAP, 40% error; in the control condition, 1 in the first IAP, 45% error). These participants did not produce extreme scores in the IAT. All other error rates were below 24%. Therefore, the IAP data of these 2 participants were excluded from all analyses. The distributions of the log-based IAP and IAP retest were not even marginally different from a normal distribution, $Z < 1$. The internal consistency of the two IAPs was evaluated similarly to the IATs by computing Cronbach's α for the separately determined IAP scores for 4 blocks of trials (3-32, 33-64, 65-96, 97-128). The overall internal consistency was .83 for the test and .77 for the retest but was somewhat unsatisfactory in the control group for the retest. In particular, it was .82 in the faking versus .86 in the control condition for the test, and .81 versus .55 for the retest, respectively. Nevertheless, internal consistency was completely satisfactory for the first test at least. The retest reliability of the IAP was $r = .65$ and thus highly similar to the retest reliability of the IAT of .68.

Effects of instruction, position, and their interaction on the IAP means were tested by a 2×2 ANOVA. No significant effects were found. In particular, the instruction effect was not even marginally significant, $F_{(1,294)} = 1.61$, $p = .21$. Thus, the IAP tended to be more robust than the IAT with regard to faking. This conclusion was also supported by an analysis of reported faking. Of the 239 participants in the faking condition, 68 reported attempts of influencing the IAP outcome. In 64 cases, they reported to have taken the perspective of a nonshy job applicant; 4 other participant reported to have deliberately committed errors. These figures were slightly higher than for the IAT. However, a t test contrasting them with the other 171 participants in the faking condition did not even reveal marginal differences, $t < 1$. In terms of untransformed reaction times, the IAP score was -91 ms for fakers, -83 ms for assumed nonfakers, and –62 ms for control participants. Although the rank-order of these means was identical with the results for the IAT, the differences between the means were minimal.

Direct self-ratings. All self-rating scales showed a satisfactory internal consistency, $\alpha > .80$. Both shyness means in the control condition were not even marginally different from those in Study 1 by Asendorpf et al. (2002), $t < 1$, which suggests that the sample of the control condition was not differently selected for shyness from the sample of this earlier study. Effects of instruction, position, and their interaction on the shyness self-ratings were tested by a mixed 2×2 ANOVA with instruction as a between-subjects factor and order as a within-subjects factor. A very large instruction effect was found, $F_{(1,298)} = 298.9$, $p < .001$. As Table 9 indicates, participants in the faking condition reported shyness that was 2 standard deviations lower than in the control condition. In addition, a moderate position effect was found, $F_{(1,298)} = 13.25$, $p < .001$, $d = .40$ (computed as $\sqrt{2}(M_1 - M_2)/SD$ where SD is the standard deviation of the difference scores; see Cohen, 1988). Participants in the faking and in the control group reported somewhat less shyness after the role play than before (see Table 9). This may be attributed to the mastery of the role play that probably made participants to consider themselves as less shy than before. The position by instruction interaction was not significant, $F < 1$.

It should be noted that position effects on direct shyness measures were analyzed in a within-subjects design whereas position effects on indirect shyness measures were analyzed between subjects (see the overall design in Table 8). Thus, comparing results for direct and indirect measures is not entirely fair, since the statistical tests had a higher level of power for the former than for the latter. However, analysis of means did not indicate any

common trend for position effects on indirect shyness measures. In terms of untransformed reaction times, the mean IAT score was -130 ms (*SD* = 214 ms) before and -85 ms (*SD* = 159 ms) after the role play. Thus, participants were more likely to attain higher shyness scores after the role play. The opposite was true for the IAP, -73 ms (*SD* = 146 ms) before and –88 (*SD* = 125 ms) after the role play. Given the *standard deviations*, though, none of these differences were significant and should be seen as chance variations. Thus, the indirect measures, in fact, seemed to be more robust against position effects.

Most other direct self-ratings showed also large instruction effects, particularly the social desirability scale, $d = 2.06$, but also the bipolar self-ratings of conscientiousness, $d = 1.34$, and intellect, $d = 1.23$. Thus, the participants in the faking condition showed a strong, generalized tendency to present themselves in socially desirable ways.

4.5.2 Effects of Instruction on Behavioral Shyness Measures

In this section, the judgments and codings of shy behavior are described, and effects of instruction (main faking effects, Hypothesis 1) are analyzed.

Judgments of shy behavior. In the control condition, the mean of the observer judgments of shyness was marginally higher than in Asendorpf et al.'s (2002) Study 1, $t_{(196)} = 1.65$, $p < .10$. Because the observers used a response scale that was anchored with extreme examples from this earlier study, this difference can be attributed to a slightly more successful induction of shyness by the role play. Table 9 indicates that the participants in the faking condition were judged as less shy than those in the control condition but this instruction effect ($d = 0.27$) was not large compared to the effect for the direct ratings.

Codings of shy behavior. The durations of the 3 types of body movement were skewly distributed and therefore log(x+1)-transformed. Table 9 indicates that, as expected, the participants in the faking condition talked more and used somewhat more illustrating gestures (significant for a one-tailed test), thus showed less controlled shy behavior than those in the control condition. In contrast, they did not show less spontaneous shy behavior with regards to facial or body adaptors, and even showed *higher* body tension, when they were instructed to appear non-shy. This behavioral pattern completely replicates the pattern that Asendorpf et al. (2002) reported for a much smaller, female-only sample. Thus, the participants in the faking condition followed the instruction to present themselves as non-

shy in their controlled behavior. However, they failed to suppress or even showed more spontaneous shy behavior than in the control condition.

4.5.3 Correlational Analyses

In the preceding analyses, I explored main effects of instruction (faking versus control) and position (before versus after the role play). In this section, I study differential effects of faking and position, that is, how faking and position affected interindividual differences and their correlates (Hypotheses 3-6).

Position effects. Explored were position effects on the correlations between the implicit and direct self-concept measures, the observer judgments, and the behavior codings, both overall and within the faking and the control group. All order effects were small and not even marginally significant. Although relatively large samples are needed to detect significant differences between correlations, the sample size for the two positions for the faking condition was $n = 120$ and thus sufficient for detecting marginally significant between-correlation differences of approximately .30 or larger with a power of .80 (Cohen, 1988). In particular, no systematic trend was found that the direct or indirect self-concept assessments before the role play were less strongly related to the behavioral observations than the same assessments after the role play. Furthermore, the self-ratings before and after the role play correlated above .83 in both the faking and the control condition, which is close to the reliability of these ratings. Therefore, the two bipolar shyness self-ratings were averaged for each participant, yielding one aggregated index of the explicit self-concept of shyness, and the position of the indirect measure was ignored in the following analyses.

Table 10 indicates that IAT and IAP were moderately correlated in both the faking and the control group and showed highly similar correlations with the other main variables. Thus, all major IAT correlates were replicated with the IAP. Therefore, both IAP and IAT were z-transformed within experimental condition to make their scores comparable, and then averaged, yielding one aggregated index of the implicit self-concept of shyness. The remaining analyses of differential effects (Hypotheses 3 - 6) were restricted to the aggregated measures of the explicit and the implicit self-concept of shyness (lower right-hand side of Table 10). Numerous observations can be made from this part of Table 10.

First, as expected by Hypothesis 3, the implicit and explicit self-concept measures were significantly less strongly correlated in the faking condition than in the control condition, $r = .19$ vs. $r = .50$, $z = 2.39$, $p < .01$, one-tailed. The correlation of .50 in the

control condition was similar to the correlation of .44 between the indirect and direct measure in Asendorpf et al.'s (2002) Study 1.

Table 10

Correlations of the Main Variables by Instruction

	1	2	3	4	5	6	7	8
1. IAT		.50***	.87***	.15*	-.07	.14*	.04	.06
2. IAP	.44***		.87***	.18**	-.09	.10	.04	.03
3. Implicit shyness[a]	.85***	.85***		.19**	-.09	.14*	.05	.05
4. Explicit shyness[b]	.35**	.49***	.50***		-.48***	.13*	-.01	.06
5. Social desirability	-.13	-.09	-.13	-.17		-.08	-.04	-.03
6. Observer judgment	.17	.28*	.27*	.36**	.16		.19**	.47***
7. Spontaneous behavior[c]	.04	.07	.07	.15	.04	.34**		.02
8. Controlled behavior[d]	.10	.02	.07	.18	.05	.70***	.29*	

Note. Correlations above the diagonal refer to faking condition (n = 238), correlations below the diagonal to control condition (n = 58). * $p < .05$ ** $p < .01$ *** $p < .001$.

[a] Mean of z-transformed IAT and IAP.

[b] Mean of the bipolar shyness self-ratings before and after the role play.

[c] Average of z-transformed duration of facial and body adaptors and tense body posture.

[d] Average of reversed z-transformed duration of speech and illustrators.

Second, as expected by Hypothesis 4, the indirect measure did not correlate with social desirability neither in the faking nor in the control group, r = -.09 and r = -.13. In contrast, the direct measure correlated significantly more negatively with the social desirability index in the faking condition than in the control condition, r = -.48 vs. r = -.17, z = 2.41, $p < .01$, one-tailed. As pointed out in the introduction, this correlational difference confirms the undermining effect of differential self-presentation tendencies on the direct shyness ratings in the faking condition.

Third, as expected by Hypothesis 5, the correlation of the observer judgments of shyness with the direct self-ratings of shyness decreased significantly under faking (from r = .36 to r = .13, z = 1.67, $p < .05$, one-tailed) whereas the correlation with the indirect measure did not (from r = .27 to r = .14, z = .92, *n.s.*, one-tailed). Although the difference

in the decrease of the correlations was not significant, it should be noted that the indirect and the direct measure showed significant and equally strong associations with the observer judgment under faking.

Because the indirect and direct measures were correlated, and to a different degree in the faking and the control group, I analyzed unique contributions of the indirect versus direct measures to the observer judgments, using multiple regression. In the control group, only the direct self-ratings explained significant unique variance of the observer judgments, $\beta = .30$, $p < .05$, whereas the unique contribution of the indirect measures was not significant, $\beta = .11$, $p = .41$. In the faking group, both measures explained similar unique but small variance that was significant for the direct self-rating, $\beta = .13$, $p < .05$, and marginal for the indirect measure, $\beta = .11$, $p < .10$. Thus, whereas the unique contribution of the direct self-ratings tended to be smaller in the faking than in the control group, the unique contribution of the indirect measure was the same in both groups. These findings suggest that the observers were to some extent resistant to participants' differential cheating. That was also indicated by the nonsignificant correlations between the observer judgments of shyness and the social desirability index in both groups (see Table 10).

The correlation of .19 ($p < .01$) between participants' direct self-ratings and the indirect measure under faking suggests that these self-ratings were not completely invalid for participants' true self-concept. This assumption was supported by a similarly high correlation of .24 ($p < .001$) between the self-ratings of shyness that were completed under the faking versus the honesty instruction at the end of the experiment. Although this correlation is much lower than the retest correlation of .83 under faking, the rank order of the participants in self-reported shyness was preserved to some extent despite differential faking. Thus, the significant but low correlation of .13 between the observer judgments and the direct self-ratings under faking may reflect this valid portion of the direct self-ratings rather than a faking effect on the observer shyness judgment.

Fourth, the indices of spontaneous and controlled shy behavior were significantly correlated with the observer judgment of shyness in both the faking and the control group. However, contrary to Hypothesis 6, both behavioral indices of shyness were not significantly correlated with the indirect and direct shyness measures in the control condition (both behavioral indices were significantly correlated with the indirect and the direct shyness measures in Asendorpf et al.'s, 2002, Study 1). Thus, although the observers interpreted these two behavioral indices as indicators of shyness, they were in fact

unrelated to the self-concept of shyness. This lack of validity applied not only to the aggregated behavioral indices but also to each single behavioral variable. Because of these zero correlations, the expected double dissociation between the indirect and direct measures of shyness was not found for the control situation.

Fifth, given this lack of validity of the behavioral measures for the control condition, it is not surprising that they lacked validity also in the faking condition. Again, the correlations between the indirect and the direct shyness measures and the two behavior composites (and each single behavior within the composites) were not significant. Therefore, the expected double dissociation between the indirect and direct measures of shyness was not found also for the faking condition.

All in all, the 4 hypotheses concerning correlations between the implicit and the explicit self-concept of shyness, the social desirability index, and the observer judgments of shyness were at least marginally confirmed but not Hypothesis 6 because of the invalidity of both the spontaneous and the controlled behavioral measures for the role play situation.

Exploration of alternative behaviors indicating shyness in the role play. The significant correlations between the observer judgments and the implicit and explicit self-concept measures suggested that the observers were aware of interindividual differences in shyness but used different cues than those captured by the *a priori* defined spontaneous and controlled behavioral measures. Therefore, alternative behavioral measures of the self-concept of shyness were systematically explored, using the videotapes of both the control condition and the Asendorpf et al. (2002) Study 1. As a safeguard against chance findings, given the *post hoc* nature of these analyses, only those behavioral measures were considered that correlated significantly with the implicit or explicit self-concept of shyness in both the control role play situation and in Asendorpf et al.'s (2002) Study 1. More than a dozen different nonverbal behaviors were explored for this purpose (e.g., body posture, facial cues, vocal cues, a detailed analysis of speech pauses of different length) but not a single behavior was found that survived this test. Thus, it seems that shyness is differently expressed in behavior in the role play situations of the present study than in the more naturalistic interactions with a confederate in Study 1 by Asendorpf et al. (2002). Nevertheless, it was not possible to identify the cues that the observers used for their valid shyness judgments.

4.6 Discussion

4.6.1 Summary of the Main Findings

Study 1 tested six hypotheses on the differential operation of indirect versus direct measures of the personality self-concept under naturalistic faking conditions. The indirect measures were an Implicit Association Test (IAT) and a newly developed Implicit Association Procedure (IAP). I discuss the results separately for each hypothesis, contrast the two indirect procedures with one another, and then briefly discuss general conclusions and open questions for the indirect assessment of personality self-concept.

As expected in *Hypothesis 1*, the direct self-ratings of shyness, the social desirability scores and the controlled shy behaviors decreased under faking; the decrease was particularly strong for the questionnaire measures (approximately 2 standard deviations). Also in line with this hypothesis, the IAP scores and the spontaneous shy behaviors did not decrease, supporting their non-fakability, and replicating Asendorpf et al.'s (2002) Study 2 findings. There was a slight tendency of the IAT scores to decrease under faking, but a more detailed analysis showed that this decrease was restricted to a minority of participants who had spontaneously attempted to vividly imagine themselves as a nonshy job applicant. Comparable effects of mental imagery on IAT scores and priming measures have been reported in studies that experimentally induced mental imagery of counter stereotypes (Blair, Ma, & Lenton, 2001). It should be noted, however, that even for these participants the effect was only moderate (less than a quarter of a standard deviation).

Whether the indirect tests were completed before or immediately after the shyness-inducing role play had no effects on their mean level, confirming *Hypothesis 2*. In contrast, the direct self-ratings of shyness were lower after the shyness-inducing role play. This may be attributed to the mastery of the role play that decreased the direct shyness self-ratings. The higher robustness of the indirect measures against state effects is important for the interpretation of the indirect measures because they are assumed to refer to a relatively stable self-concept of personality, not to current states (cf. Schmukle & Egloff, 2003).

Turning to the correlational hypotheses, the implicit and explicit self-concept measures were significantly less strongly correlated in the faking condition than in the control condition, which fully confirmed Hypothesis 3. This hypothesis was based on the assumption that the direct self-ratings were less valid in the faking condition than in the

control condition because they were distorted by differential tendencies of the participants to present themselves in socially desirable ways.

This assumption was supported by the finding that the indirect measures did not correlate with participants' social desirability scores under both experimental conditions whereas the direct self-ratings correlated more negatively with the social desirability scores in the faking condition than in the control condition. Thus, *Hypothesis 4* was fully confirmed. Together, these findings for the effect of faking on the correlations between the indirect measures of shyness, the direct self-ratings of shyness, and social desirability scores strongly support the view that the indirect measures were robust with regard to interindividual differences in faking attempts.

Hypothesis 5 on the validity of the observer judgments of shyness was marginally confirmed. Whereas the observer judgments tended to correlate more strongly with the explicit than with the implicit self-concept of shyness in the control group, these correlations were virtually identical under faking. Moreover, the correlation between the indirect measures and the observer judgment was similar under both experimental conditions. The unique contribution of the direct self-rating under faking to the observer judgments, independent of the contribution of the indirect measure, does not necessarily indicate that the observers were influenced by participants' faking attempts because there were two indications that participants' true shyness perspired in their behavior in the faking condition: a significant correlation for the direct self-ratings of shyness between the faking and the honesty condition, and a significant correlation between the direct and indirect measure under faking.

Together, these results suggest that observer judgments of temperamental traits in role play situations are not very much influenced by the role players' self-presentation even when they systematically try to fake the cues that the observers might use for their judgments (in this case, cues for non-shyness such as talking and gesturing). It seems that the observers use other cues that the participants cannot easily control. Unfortunately, it was not possible to identify such cues from the videotaped behavior.

Turning finally to *Hypothesis 6* on a double dissociation between indirect and direct measures, the observer judgments of shyness correlated significantly with both the spontaneous and the controlled indices of shy behavior under both experimental conditions. This validated the spontaneous and controlled indices as behavioral measures of shyness. However, these correlations were smaller than in Asendorpf et al.'s (2002)

Study 1, and contrary to these prior findings, controlled shy behavior was not significantly correlated with the direct shyness self-ratings, and spontaneous shy behavior was not significantly correlated with the indirect measures of shyness. Thus, although the observers interpreted these two behavioral indices as indicators of shyness, they were in fact unrelated to the self-concept of shyness in the control condition. Therefore, the expected double dissociation between the indirect and direct measures of shyness was not found for the control condition.

Because the mean direct self-ratings of shyness in the control condition were not lower than in Asendorpf et al.'s (2002) Study 1, and the observers rated the participants in the control condition even slightly more shy than the participants in this earlier study, the lack of validity of the behavioral measures cannot be attributed to an insufficient induction of shyness by the role play procedure. Instead, it seems that the role play framework itself, the more structured situation (a clear communication goal was defined) and/or the clear status differences between the participants ("boss" versus "employee") apparently changed the meaning of behaviors that were found to be valid indicators of shyness in an unstructured interaction between strangers.

Given this lack of validity of the behavioral measures for the control condition, it was not surprising that all correlations between the behavioral measures and the indirect and direct measures of the self-concept of shyness were not significant also in the faking condition. Therefore, the expected double dissociation between the indirect and direct measures of shyness was again not found.

Thus, in my view the main problem of the present study did not concern the indirect procedures. Instead, it concerned the fact that valid behavioral cues for shyness in more naturalistic situations became completely invalid in a role play context, and could not be replaced by alternative valid cues. If the assessment of shy behavior in role play situations is not the focal point, as in the present study, future studies might try to circumvent this problem by motivating participants to fake non-shyness in dyadic interactions of the type used by Asendorpf et al. (2002), Study 1. I did not follow these lines because I feared that direct instructions to do so would be perceived by the participants as artificial and would therefore insufficiently bias their actual behavior. Alternatively, it might be possible to motivate participants more indirectly to fake non-shyness.

4.6.2 An Alternative Procedure: The IAP

The Implicit Association Procedure (IAP) produced results that were highly similar to those found for the Implicit Association Test (IAT). The error response rate was similar to the IAT, the distribution of the scores was also close to a normal distribution, and the retest correlation was virtually identical. The internal consistency was slightly higher for the IAP which can be attributed to the fact that there were 256 trials in the critical blocks in the IAP, but only 160 trials in the IAT. The total test durations were not different, though, because there is no need in the IAP for attribute and reverse target discriminations. The two indirect tests showed substantial correlations of .50 (faking condition) and .44 (control condition), and their correlations with external variables were highly similar. A minor difference was that the IAP tended to be slightly more robust against faking.

A disadvantage of the IAP is that it is more difficult to implement this procedure on standard computers than the IAT. A joystick is needed, the joystick has to be continuously calibrated, and the program routines for implementing the procedure are much more complex than for the IAT. All in all, then, the IAP may be considered less an alternative to the IAT than an useful addition to the IAT that allows one to replicate IAT-findings and to reduce method variance of the IAT by aggregating IAT and IAP scores.

Let me conclude with a comment on the utility of indirect measures for the assessment of personality differences. On the positive side, the study showed that these indirect measures were fairly robust to faking attempts of the participants. Only participants who tried to bias their results by deliberately taking the perspective of a non-shy person were able to bias their IAT scores (but not their IAP scores), and this bias was very small compared to the bias in their direct self-ratings. Also, it was possible to construct a new indirect assessment procedure, the IAP, which correlated .50 with the IAT and showed highly similar correlates. Between-procedure correlations of this size are rarely achieved for indirect procedures that assess the same construct (Bosson et al., 2000; Cunningham et al., 2001). This new method made it possible to increase the reliability and validity of the assessment of the implicit self-concept through the aggregation of both procedures.

On the negative side, the direct self-ratings predicted the observer judgments in the control condition slightly better than the indirect measures, and in the faking condition not worse than the indirect measures. Although the direct self-ratings were strongly biased

with regard to their mean, there were multiple indications that they were not completely invalid with regard to their interindividual differences. Furthermore, the .50 correlation between the IAT and IAP is not high compared to the .70 correlations that are regularly achieved when the same personality dimension is self-rated on different questionnaire scales. Both the relatively low retest correlation of .65 - .68 for the IAP and IAT and their relatively low parallel test reliability of .50 indicate that the amount of specific method variance for these indirect procedures is much higher than the specific method variance for direct ratings.

Much work may be required to increase these methodological weaknesses of the current indirect procedures for the assessment of stable interindividual differences to a psychometrically satisfactory level. Unless such a satisfactory level is reached, the indirect procedures can be considered interesting research instruments in need for improvement, not methods that are ready to be applied for practical assessment purposes. Another important aspect concerning practical implications is whether indirect measures, similar to direct measures, allow for the concurrent assessment of more than one personality trait. Therefore, the next study explores whether the IAT may be used to assess two traits, anxiousness and angriness, within one sample.

5 Study 2: Concurrent Assessment of the Implicit Self-Concept of Anxiousness and Angriness

5.1 Introduction

The results of Study 1 showed that indirect measures (shyness IAT and IAP) do not yet meet psychometric criteria that are necessary for individual diagnosis and that are typically shown for direct measures, that is, satisfactory test-retest stability and high convergent validity. The main purpose of Study 2 was to examine another important aspect concerning the practical implications of indirect assessment. Direct self-reports, for example, the NEO-PI-R (Costa & McCrae, 1992), allow for the concurrent assessment of different traits within one questionnaire. Therefore, Study 2 explored whether IATs also allow for the assessment of two different traits within one sample when the IATs are applied as two consecutive tests. Although several studies employed more than one IAT within one sample (e.g., Gawronski, 2002; Nosek, Banaji, & Greenwald, 2002), there appears to be no research that would systematically carry out position effects on the IAT.

Therefore, the sequence of an anxiousness and an angriness IAT was counterbalanced across participants in Study 2 and three main research questions were explored. First, it was expected that the validity of the IAT is affected if the IAT is preceded by another IAT. Second, it was expected that the IATs add incremental validity to the prediction of anxious and angry behavior. Third, it was explored whether social desirability does moderate the relationship between direct and indirect measures. These research questions are discussed in more detail in the following sections.

5.1.1 Research Question 1: Position Effects on IATs

The sequence of IATs is often counterbalanced in studies that explore several IATs. Usually, results are not discussed separately for the groups of different IAT order. Concerning test-retest comparisons, the study by Asendorpf et al. (2002) provided evidence that the second, parallel shyness IAT tended to show lower correlations with direct shyness measures and with shy behavior than the first IAT. Similarly, other studies found that the retest reliabilities of IATs are lower than their internal consistencies (cf. for an overview, Egloff, Schwerdtfeger, & Schmukle, 2003). Thus, IAT measures showed both a validity decrease for the second test and relatively low test-retest reliabilities. Both

aspects might be caused by the two factors, that is, state influences and changes in response strategies.

State influences were ruled out in Study 1 and in the study by Schmukle and Egloff (2003) as a systematic bias in IAT results. Therefore, it is most likely that differences between the first and the second IAT measures are due to changes in response strategies. Working on the IAT, participants might develop cognitive strategies to respond faster, and try out different response styles, for example, avoiding errors because errors increase test duration. Recently, De Houwer (2003a) stated that changes in response strategies may emerge because participants try to make the IAT tasks as simple as possible. Therefore, participants *recode* the double discrimination task in terms of a simple discrimination (e.g., positive versus negative; see Mierke & Klauer, 2001). De Houwer pointed out that recoding in terms of a simple discrimination may be based on the associations one aims to measure (e.g., the associations of flowers and insects with positive and negative attributes). Alternatively, recoding may be based on any type of similarity between target and attribute concept (e.g., word length, color, etc.). Importantly, IAT effects are likely to be distorted if the similarity-based task-recoding is unrelated to the associations one tries to assess.

Study 2 examined whether the completion of an IAT distorts the validity of the succeeding IAT due to any change in response strategies. Therefore, the order of the anxiousness and the angriness IAT was counterbalanced in Study 2, and results were inspected separately for both groups of different IAT order.

5.1.2 Research Question 2: Prediction of Anxious and Angry Behavior

Recently, a study by Egloff and Schmukle (2002) showed that self-reported state anxiety during a stressful speech was predicted by direct anxiousness measures but not by the anxiousness IAT. More importantly, the observer judgments of anxious behavior and several behavioral indicators of anxiety were predicted by the anxiousness IAT but not by direct anxiousness measures. Using the same rationale, Study 2 examined whether indirect measures significantly increase the prediction of behavior even if two traits are assessed within one study. It was expected that both, the anxiousness and the angriness IAT show predictive validity for anxious and angry behavior, respectively.

Anxiousness and angriness were employed as traits under investigation because they were expected to be uncorrelated at least at the level of direct self-reports. Uncorrelated traits facilitate the study of convergent and discriminant validity between

direct, indirect, and behavioral measures (see Chapter 2.5 and the following). Additionally, anxiousness and angriness allow for the study of the predictive validity of direct and indirect measures because anxious and angry behavior may be observed after emotion inductions.

5.1.3 Research Question 3: Social Desirability as a Moderator Variable

One of the main reasons for research interest in indirect measures is that they are expected to circumvent the validity problems that are associated with direct measures (Greenwald et al., 2002). An example of a validity problem in direct measures is their susceptibility to social desirability concerns . For example, it was shown in Study 1 and in other studies (e.g., Asendorpf et al., 2002; Egloff & Schmukle, 2002) that direct self-report measures were, in contrast to IAT measures, significantly correlated with social desirability. Social desirability is a tendency to portray oneself in a favorable light (Crowne & Marlowe, 1960). Therefore, the more negative the correlations between direct measures and social desirability are, the more biased by social desirability the direct measures are assumed to be.

More importantly, social desirability may also act as a moderator variable in the relationship between direct self-report and IAT measures. Individuals with a weak tendency to present themselves in a socially desirable way should show higher correlations between direct measures and IATs than the individuals with a strong tendency to socially desirable responding. Previous studies indicated that the correlations between direct measures and IATs were not moderated by social desirability (Egloff & Schmukle, 2003; Hofmann, Gschwendner, & Schmitt, 2003). In contrast, moderator variables were found to be significant if they asked for self-presentational motivation more directly with regard to the attribute that was measured (e.g., Banse & Gawronski, 2003; Hofmann, Gschwendner, et al., 2003; Nosek & Banaji, 2002). Nevertheless, social desirability was explored as a moderator variable in Study 2 in order to replicate the results from other studies for the anxiousness and the angriness IAT.

5.2 Hypotheses

Study 2 tested the following hypotheses.

Hypothesis 1 (Increase of state anxiety and state anger). Participants report more state anxiety and state anger after the emotion inductions as compared to the baseline.

Hypothesis 2 (Validity of the bipolar anxiousness and angriness self-ratings). The bipolar anxiousness self-rating correlates with direct anxiousness but not with direct angriness measures whereas the opposite is true for the bipolar angriness self-rating. This validates the word material that was used in the IATs.

Hypothesis 3 (Zero correlation between social desirability and the IATs). In contrast to direct self-ratings neither the anxiousness nor the angriness IAT are correlated with social desirability scores.

Hypothesis 4 (Social desirability is not a moderator variable). Social desirability does not moderate the correlations between indirect and direct measures.

Hypothesis 5 (Zero correlation between anxiousness and angriness). Anxiousness and angriness are neither correlated for the direct, nor the indirect or the behavioral measures, confirming their conceptualization as orthogonal dimensions.

Hypothesis 6 (Validity decrease for the second IAT). The IAT tends to show smaller convergent validity with direct and behavioral measures when it is preceded by another IAT.

Hypothesis 7 (Independent contribution of IATs to behavior prediction). The anxiousness and the angriness IAT predict behavioral anxiety and anger even when direct self-ratings are controlled for. In contrast, self-reported state anxiety and state anger are predicted by direct self-ratings but not by the anxiousness and angriness IAT.

5.3 Methods

5.3.1 Participants

A total of 103 university students were recruited as participants on the campus of Humboldt University, Berlin, none of whom were psychology students or had participated in the lab's previous studies. Most participants were directly approached by an experimenter (not identical with the experimenter at the lab). Some participants were recruited using postings at the university buildings. Participants were asked to take part in a study on concentration and personality. As a compensation, participants were offered €

10 (approximately US $ 10 at the time) for completing a questionnaire of about 15 minutes duration at home and for participating in a lab experiment of about one hour duration. In addition, they could receive individual feedback on their results after the study is complete. All participants claimed to be native German speakers. Three female participants refused to complete the speaking task during the lab session, and were therefore excluded from analysis. This led to a final sample of 100 participants (50 male, 50 female; age *M* = 24.0 years, range 19-32 years).

5.3.2 Assessments and Measures

Overall procedure. The overall procedure of Study 2 is depicted in Table 11. All participants (a) judged themselves on several trait measures at home within one week before the lab session. After arrival at the lab they (b) completed a short form of the d2 Attention-Stress Test, (c) judged themselves on a short optimistic risk perception measure (not relevant to this research), (d) completed the anxiousness IAT and the angriness IAT, (e) indicated their state anxiety and state anger on bipolar items, (f) received instructions for an anxiety-inducing speech, (g) completed a retest of (e), (h) prepared their speech, (i) were video-taped during their speech, (j) were videotaped during an anger-inducing computer crash, (k) completed a retest of (e), (l) were interviewed about the experiment, and (m) were completely debriefed.

The anxiousness and angriness items of the two IATs were included as direct self-ratings in step (a), (e), (g), and (j). The order of the anxiousness IAT and the angriness IAT in step (d) was varied between participants such that half of the participants completed the anxiousness IAT first and the other half completed the angriness IAT first. The assignment to the two orders was balanced for gender and alternated between successive participants. In contrast, the order of the anxiety and the anger induction was fixed, such that the anxiety induction always came first, because of the faked computer crash during the anger induction.

Finally, the participants were thanked and asked to give their consent for the evaluation of the videotapes (all agreed). They were also paid and promised individual feedback about their results. Four months after finishing data collection, participants received a letter containing the principal findings of the study along with an invitation for an individual feedback session, where they were informed about their personal results.

Table 11

Overall Procedure of Study 2

Cover story: Concentration and personality	Duration (Min.)
At home:	
(a) Direct trait measures - Trait form of the STAI, STAXI, and two subscales of the TAI-G - Speaking Anxiety Scale - Bipolar self-ratings of anxiousness, angriness, conscientiousness, and intellect - Social desirability scales and MAS - Biographical data	15
At the lab:	
(b) d2 Attention-Stress Test	5
(c) Optimistic risk perception measure	2
(d) Anxiousness and angriness IAT (counterbalanced for order across participants)	20
(e) Direct state measures (baseline) Bipolar self-ratings of anxiety, anger, and conscientiousness	1
(f) Anxiety induction: Instructions for the speech	2
(g) Direct state measures Bipolar self-ratings of anxiety, anger, and conscientiousness	1
(h) Preparation of the speech	3
(i) Behavior observation: Speech before video camera	5
(j) Anger induction and behavior observation: Computer crash, which was (1) pretended to be caused by the participant (2) destroyed all his / her data (3) made payment of the reward impossible	5
(k) Direct state measures Bipolar self-ratings of anxiety, anger, and conscientiousness	1
(l) Interview: Identification of participants who doubted the computer crash	5
(m) Debriefing about the true purpose of the study	5
	70

Note. STAI = State Trait Anxiety Inventory, STAXI = State Trait Anger Expression Inventory, TAI-G = German version of the Test Anxiety Inventory, MAS = Manifest Anxiety Scale, IAT = Implicit Association Test.

Trait measures. In order to minimize transfer effects between direct and indirect measures, direct trait measures were mailed to the participants within one week before the lab session. The instructions explained to participants that the study was about concentration and personality and consisted of two parts: a set of questionnaires concerning several personality traits, that was attached and had to be completed at home, and a subsequent lab session assessing attention and concentration. I avoided to tell participants that the study was about anxiousness and angriness because I (a) did not want anxious persons to avoid participation in the study, and (b) had to keep participants naive about the anger induction, as most people would not get angry knowing that it was intended to provoke their anger (Stemmler, Heldmann, Pauls, & Scherer, 2001).

The mailed questionnaire contained the following measures (test references list the used German version first, and the English equivalent second, if such equivalent existed). The questionnaire started with the trait forms of the *State Trait Anxiety Inventory STAI* (Laux, Glanzmann, Schaffner, & Spielberger, 1981; Spielberger, Grousch, & Lushene 1970) and the *State Trait Anger Expression Inventory STAXI* (Schwenkmezger, Hodapp, & Spielberger, 1991; Spielberger, 1988) together with the subscales Interference and Lack of Confidence (without the item "Ich bin überzeugt, dass ich gut abschneiden werde." ["I am sure, that I will receive good marks."]) of the *Test Anxiety Inventory TAI-G* (Hodapp, 1991; expanded *G*erman version of the TAI, Spielberger, 1980). These questionnaires assess enduring symptoms of anxiousness, angriness, and test anxiousness on a 4-point scale (1 = Almost never, 4 = Almost always) with 20, 10, and 11 items, respectively. The TAI-G subscales were added, and all scales were mixed in a fixed random order, because participants of a pilot study doubted the cover story when the STAI and the STAXI were presented separately. When both scales were mixed with the TAI-G, the STAI and the STAXI were less salient cues for the true content of the experiment.

The trait measures proceeded with the second series of the *Speaking Anxiety Scale* (Spitznagel, Schlutt, and Schmidt-Atzert, 2000). This questionnaire assesses habitual emotionality (e.g., "I am quite nervous") and worries (e.g., "I fear negative consequences") immediately before giving a speech with 8 items each. Items were presented on a 4-point scale (1 = I do not agree at all, 4 = I agree completely).

Subsequently, participants had to rate their conscientiousness, intellect, attentiveness, anxiousness, and angriness on *33 bipolar adjective pairs* (e.g., " self-confident 1-2-3-4-5-6-7 anxious"). The pairs were mixed in a fixed random order and

presented with a trait instruction. The *10 intellect* and *10 conscientiousness* pairs were the same as in Asendorpf et al.'s Study 1 (2002). I further added *3 attentiveness* pairs to make the cover story more plausible. The first pair was "aufmerksam" ["attentive"] versus "durcheinander" ["jittery"] that was adapted from the Positive and Negative Affect Schedule PANAS (Krohne, Egloff, Kohlmann, & Tausch, 1996; Watson, Clark, & Tellegen, 1988). Two additional pairs were synonymous.

The *5 anxiousness* pairs (anxious versus self-confident) and the *5 angriness* pairs (angry versus self-controlled) were constructed on the basis of 430 unipolar and 179 bipolar adjective items provided by Ostendorf (1990). He had factor analyzed them and reported their loadings on the first five factors that could be interpreted as the factors of the Five Factor Model of personality. Within the Five Factor Model (see Chapter 2.6.2), anxious versus self-confident was conceptualized as being strongly related to neuroticism, moderately related to introversion, and as being unrelated to agreeableness. In contrast, angry versus self-controlled was conceptualized as being weakly related to neuroticism and extraversion, but as being strongly negatively related to agreeableness.

Consequently, concerning the anxious pole, I selected unipolar items with factor loadings above .25 on both introversion and neuroticism, and below .10 on agreeableness. For the opposite pole, self-confident, unipolar items representing the inverse factor loadings were selected. Concerning the angry pole, I selected unipolar items with factor loadings above .20 on extraversion and neuroticism, and below -.25 on agreeableness. For the opposite pole, self-controlled, unipolar items representing the inverse factor loadings were selected. 9 items met these requirements. Then, I searched for bipolar adjective pairs that showed the same pattern of factor loadings, and received another 13 adjectives. Finally, I added 6 self-generated, semantically similar adjectives. This procedure resulted in a list of 7 bipolar items describing anxious versus self-confident, and 7 bipolar items describing angry versus self-controlled. These items were pre-tested in a student sample ($N = 42$; age $M = 22.6$ years, range 19-39 years) together with three scales of seven bipolar adjectives from the 179 items list, which had the highest factor loadings on either the neuroticism, the extraversion, or the agreeableness factor and cross-loadings below .30. Within the 7 anxiousness pairs, 5 showed significant negative correlations with extraversion ($r < -.32$ $p < .05$); the two noncorrelating items were excluded. The resulting 5 item bipolar anxiousness scale showed acceptable internal consistency, $\alpha = .84$, and correlated strongly with neuroticism ($r = .82$; $p < .001$), intermediately with extraversion

(r = -.45; p = .003), and nonsignificantly with agreeableness (r = -.19). From the seven angriness pairs, two pairs that showed significant positive correlations with the anxiousness scale were excluded. The resulting 5 item bipolar angriness scale showed acceptable internal consistency, α = .77 and correlated marginally with neuroticism (r = .21; p = .18) and extraversion (r = .22; p = .17), highly with agreeableness (r = -.78; p < .001), and was not correlated with the 5 item anxiousness scale (r = .01). All items of the bipolar anxiousness and angriness scale were used as word material within the IATs and are listed in Table 12.

Finally, the questionnaire concluded with the *social desirability scales* by Lück and Timaeus (1969; Crowne & Marlowe, 1960) and Stöber (1999; without the Item "Have you ever consumed drugs"). These scales contain 16 and 23 items, respectively, and measure socially desirable responding by asking for socially desirable but infrequent or socially undesirable but frequent behaviors on a true-false format. Items of both scales were presented in a fixed random order together with the *Manifest Anxiety Scale MAS* (Lück & Timaeus, 1969; Taylor, 1953). The 23 items of this scale assess various symptoms of anxiousness (e.g., "I work under a great deal of tension"). The reliability of all trait measures was satisfactory and is reported in Table 13 of the Results section.

After answering these personality items, participants had to report their age, sex, height, dominant hand, academic subject, length of time spent at university, whether they were still students (all were), and whether they had a permanent partner.

Lab session. Upon arrival at the lab participants were reminded that the experiment was about attention and concentration. The experimenter briefly explained that the lab session contained different concentration tests, two of which were on the computer, and one being a paper-and-pencil test, as well as a situation demanding attention and concentration that would be videotaped. Subsequently, participants received instructions for the first concentration test. Because men might repress their anger facing a woman, and women might avoid getting angry with a physically superior man, the experimenter was always of the same gender as the participant.

d2 Test. Instructions and procedures of the d2 Test (Attention-Stress Test, Brickenkamp, 1994) corresponded to the test manual but I only presented the first 5 rows instead of the complete 14 rows version. During the d2 Test participants are given 20 seconds per row with 47 stimuli each to cross out relevant stimuli (letter "d" with exactly two lines) and ignore irrelevant stimuli (letter "d" with more or less than two lines and any

letter “p”). The test score is calculated as the difference between processed stimuli and errors (false alarms and misses). The d2 Test was primarily used to give a better justification for the cover story. Therefore, results for the d2 Test will not be reported here.

Optimistic risk perception measure. After the d2 Test and before the IATs I presented a German translation of the optimistic risk perception measure from Lerner and Keltner (2001, Study 4) as a short break from concentration tasks. The questionnaire was presented on the computer and was added for the purpose of another study. The internal consistency of this 15-item questionnaire was low, $\alpha = .58$.

Anxiousness and angriness IAT. The procedures for the anxiousness and the angriness Implicit Association Test (IAT) were identical to the shyness IAT in Study 1. Consequently, both IATs were the same except for the attribute dimension, being anxious versus self-confident within the anxiousness IAT, and angry versus self-controlled within the angriness IAT. Task sequence and stimuli are depicted in Table 12. IAT scores were computed as D measures with an error penalty of 600 ms, and without the exclusion of trials below 400 ms (for details on the complete algorithm, see Greenwald et al., 2003). Like conventional scores, D measures were based on the difference between mean response latencies in sequence 5 and sequence 3 (see Table 12), but were scaled in units of the individuals’ standard deviations and included an error penalty for incorrect responses. In contrast to Greenwald et al. (2003), all trials were considered equally and the first 20 trials were not weighed as more important as the succeeding trails, because I (a) did not declare the first 20 trials as training trials and (b) had 60 instead of 40 succeeding trials.[1] The measures were coded so that high scores represented quicker associations of Me-anxious and Others-self-confident relatively to Me-self-confident and Others-anxious, or of Me-angry and Others-self-controlled relatively to Me-self-controlled and Others-angry, respectively. Internal consistencies are reported in the Results section.

[1] Different procedures of weighing the first 20 trials more than the succeeding trials did only minimally change the results.

Table 12

Implicit Association Tests for Anxiousness and Angriness: Task Sequence and Stimuli

			Response key assignment	
Sequence	N of trials	Task	Left key	Right key
1	40	Target discrimination	Me	Others
2	40	Attribute discrimination	Anxious (angry)	Self-confident (self-controlled)
3	80	Initial combined task	Me, anxious (angry)	Others, self-confident (self-controlled)
4	40	Reversed target discrimination	Others	Me
5	80	Reversed combined task	Others, anxious (angry)	Me, self-confident (self-controlled)

Stimuli					
		Anxiousness IAT		Angriness IAT	
Me	Others	Anxious	Self-confident	Angry	Self-controlled
I	they	anxious	self-confident	angry	self-controlled
self	them	timid	daring	hot-tempered	thoughtful
My	your	insecure	secure	undercontrolled	self-disciplined
Me	you	worried	unconcerned	hot-headed	adaptable
Own	other	overly cautious	carefree	irritable	calm

Note. The procedures of the anxiousness and the angriness IAT were identical. Words in parentheses refer to the task sequence within the angriness IAT. The original German single word stimuli are listed in the appendix.

State measures. As a manipulation check for the emotion inductions I used bipolar items for anxiousness and angriness together with a *state* instruction. These items were mixed in a fixed random order with 3 attentiveness and 7 out of 10 conscientiousness items. The items were presented in a paper-pencil version, and were identical to those completed as a trait measure at home. 3 conscientiousness items were dropped, because I expected them not to match the state instruction (e.g., “fleißig” [“industrious”] versus “faul” [“lazy”]). State measures were presented after the IATs (baseline), the instructions for the speech (anxiety induction), and after the computer crash (anger induction). Reliabilities for the state measures were satisfactory, internal consistencies were for the

anxiety scale $\alpha = .89$, for the change in anxiety (speech minus baseline) $\alpha = .78$, for the anger scale $\alpha = .80$, and for the change in anger (computer crash minus baseline) $\alpha = .74$.

Anxiety induction. Participants received instructions for the speech on a piece of paper. The paper informed participants that they should give a speech that would be videotaped and later on analyzed by experts. The duration of the speech was asked to be five minutes. Directly after this announcement participants completed the state measures. Subsequently, they were told about the subject of the speech (terminal illness and euthanasia: immoral or humane; adapted from Schmukle & Egloff, 2003) and were given three minutes for preparation. Participants were allowed to make notes during preparation, but the speech was supposed to be given without notes. Then, participants gave their speech directly in front of the video camera that was operated by the experimenter from a nearby room. Exactly after five minutes the experimenter thanked the participants and informed them that this was enough. When participants stopped talking before the five minutes were over, the experimenter prompted them to continue talking until full five minutes were up. The time period before participants continued their speech was defined as missing. For the judgments and codings of anxious behavior secondary tapes were prepared that contained the first three minutes of noninterrupted speech. The speech task was followed by the anger induction.

Anger induction. The general procedure was adapted from Wiedig (2003) and is similar to a procedure used by Bargh et al. (1996, Experiment 3). Participants completed a STROOP-Test on the computer. Again, participants were videotaped and were told that this was to evaluate their eye-blink-rate as an indicator of concentration. In fact, this was to give good reason for videotaping the interaction with the experimenter. Three minutes after starting the STROOP, the screen froze and the words “FATAL ERROR” appeared in the center of the screen. In addition, a short but intensive error sound was given, whenever a key was pressed. The experimenter, then, approached the participant and pretended to be astonished by the accident. The subsequent interaction between experimenter and participant comprised 3 different provocations. First, the experimenter accused the participant of causing the crash by incorrectly using the enter-key. Second, she or he said that all computer-based data of the participant were now destroyed. Third, due to loss of data participants could not receive any money for the experiment. After this, participants were asked to complete the state measures, waiting for a computer expert who may help to save the data. For the judgments and codings of angry behavior secondary tapes were

prepared. The recording started when the computer crashed and ended when participants began completing the state measures. For the anger judgments, a three second blue screen interval was inserted after the end of each of the 3 provocations to enable separate ratings for each provocation.

Interview. The aim of the interview was to identify participants who doubted the computer crash. Participants were first asked whether they had difficulties with any part of the experiment. Afterwards they had to say whether they noticed anything remarkable during the experiment and whether anything in the experiment seemed strange to them. All participants (11 females and 12 males) who were suspicious about the computer crash being part of the experiment were excluded from the analysis of the anger induction. These participants did not differ significantly from the remaining participants on any of the anxiousness and angriness measures.

Debriefing. Finally, participants were completely and thoroughly debriefed about the true purpose of the study. It was made sure that participants had an opportunity to relax after the disturbing computer crash, and would not leave the lab angry or upset. In the beginning of debriefing, the participants were offered some sweets by the experimenter as a compensation for a rather harsh preceding interaction. Then, participants were informed that the study was not on concentration and attention but on anxiousness and angriness, and aimed to validate new computer based measures for these traits. Thereby, the experimenter went through the crucial parts of the study (direct and indirect measures, emotion inductions) and explained why these procedures were designed to assess anxiousness and angriness. In order to keep the true purpose of the study undisclosed for the subsequent participants, the experimenter asked the participants to keep the information about the study confidential until they would receive a letter from the experimenter. This letter was sent out four months after finishing data collection and comprised the main findings of the study together with an invitation for an individual feedback session.

Judgments of anxious and angry behavior. Four student judges that were unfamiliar with the participants and blind to their data independently rated their overall impression of the participants' anxiety and anger on 7-point scales. On these scales, 7 was labeled "very anxious" or "very angry" and 1 was labeled "not at all anxious" or "not at all angry". For the anxiety judgment each minute of the 3-minute speech was judged. For the anger judgment each of the 3 provocations after the computer crash (alleged misuse of the enter key, loss of data, no money) was judged separately. This resulted in 12 anxiety and

12 anger judgments for all participants that were averaged in each case. Similarly, two judges independently rated the anxiety and the anger within the participants' voices with three ratings per scene but without watching participants. This resulted in 6 anxious and 6 angry voice judgments that were averaged in each case. The anxiety judgments (both overall and voice judgments) were anchored by a female and a male example of extremely anxious and extremely nonanxious participants from the study by Egloff and Schmukle (2002). In the same way, the anger judgments were anchored by extremely angry and nonangry examples from a study by Wiedig (2003). Interrater reliability was satisfactory for all judgments (see Results section).

Codings of anxious behavior. All Codings were done on a PC using the Computer Aided Observation System (CAOS) software. This program synchronizes video player and PC, and registers onset and offset of behavioral codings when the appropriate key is pressed. Anxiety codings were carried out for body movements and nervous mouth movements. Following Ekman and Friesen's (1972) classification, body movements were coded as illustrators (movements illustrating speech), facial adaptors (self-stimulations of the face), and body adaptors (self-stimulations of the body). For data analysis body movements were considered in terms of their relative duration of the 3 minute speech. Nervous mouth movements were coded according to Egloff and Schmukle (2002) defined as lip biting, lip licking, twitches of the mouth, and pressing of the lips. As the nervous mouth movements were short and discrete events, Egloff and Schmukle (2002) examined their frequency rather than their duration. I also considered their frequency, because their duration might in this case be overly confounded by the noise in the coders' reaction time during the on-off coding. In order to control cross-lab reliability with the Egloff and Schmukle (2002) coding system, one coder first coded 10 female and 10 male participants of Egloff and Schmukle's Study 4. This coder correlated highly with the mean of two coders of Egloff and Schmukle's (2002) study and showed therefore substantial agreement between the coding in both labs, $r = .80$. In addition, within-lab reliability of all anxiety codings was assessed by a second coder with independent codings of 20 randomly selected participants and was satisfactory in all cases (see Results section).

Codings of angry behavior. Anger codings were completed for emotional facial expressions of the Ekman and Friesen's (1978) coding system that were shown to co-occur with anger (Friesen & Ekman, 1984). These were the Action Units brow lower (AU 4), upper lip raise (AU 10), lip funnel (AU 22), lip tight (AU 23), lip press (AU 24), and chin

raise (AU 17), that were coded in independent runs. As the coded facial expressions were short and discrete events, I further considered their frequency per minute rather than their relative duration of observation time. Reliability estimates were provided by independent codings of 20 participants by another coder. Reliability was not satisfactory for the AU 10, 22, and 17, because they occurred so rarely (mean frequency less than 0.25 times per minute) that intercoder reliability was hard to obtain. Therefore, these Action Units were not considered for data analyses. Reliability for the other three Action Units was acceptable (see Results section).

5.4 Results

5.4.1 Efficacy of Emotion Inductions

To investigate whether the speech and the computer crash were apt to observe anxious and angry behavior, I first needed to examine the efficacy of these emotion inductions. As expected by *Hypothesis 1*, participants reported more state anxiety after the announcement of the speech (M = 3.39) than at the beginning of the experiment (M = 3.02), $t_{(99)} = 4.11$, $p < .001$, $d = .58$. Similarly, participants reported more state anger after the computer crash (M = 2.53) than at the beginning of the experiment (M = 2.35), $t_{(76)} = 1.98$, $p < .05$, $d = .31$. It should be noted that the degrees of freedom were smaller for the anger induction because I had to exclude participants who were suspicious about the computer crash. Considering the increase in self-reported state measures, both of the emotion inductions worked.

5.4.2 Validity of the Bipolar Anxiousness and Angriness Self-Ratings and the IAT Stimuli

This section inspects the convergent and discriminant validity of the bipolar anxiousness and angriness self-ratings that were also used as word material IATs. The reliabilities and correlations of all direct trait measures are depicted in Table 13. Reliability (Cronbach's α) was satisfactory for all measures, in particular it was .84 for the bipolar anxiousness and .80 for the bipolar angriness self-rating. As one can see in the first two rows of Table 13, the bipolar *anxiousness* self-rating correlated highly with the Manifest Anxiety Scale and the trait form of the State Trait Anxiety Inventory, and intermediately with the subscales of the Speaking Anxiety Scale. These subscales assess habitual emotionality and worries immediately before giving a speech and, in contrast to general

anxiousness questionnaires, are more situation-specific. The bipolar anxiousness self-rating also showed a small correlation with the trait form of the State Trait Anger Expression Inventory.

Table 13

Correlations between the Trait Measures in Study 2

	1	2	3	4	5	6	7	8
1. Bipolar anxiousness self-rating	*.84*	-.08	.30**	.35***	.72***	.73***	.23*	-.08
2. Bipolar angriness self-rating		*.80*	-.05	-.05	.12	.07	.45***	-.30**
3. Speaking Anxiety Emotionality			*.88*	.72***	.36***	.28**	.13	-.16
4. Speaking Anxiety Worries				*.84*	.44***	.40***	.23*	-.23*
5. Manifest Anxiety Scale					*.82*	.78***	.39***	-.30**
6. State Trait Anxiety Inventory[a]						*.90*	.37***	-.25*
7. State Trait Anger Expression Inventory[a]							*.78*	-.34***
8. Social Desirability								*.81*

Note. $N = 100$. Internal consistencies (Cronbach's α) are printed in *italics* along the diagonal.
[a] Trait form. $^{+}p < .05$ $^{*}p < .05$ $^{**}p < .01$ $^{***}p < .001$.

In contrast, the bipolar *angriness* self-rating did not even marginally correlate with any direct anxiousness measure and correlated intermediately with the trait form of the State Trait Anger Expression Inventory. Thus, the correlation for the angriness self-rating with the corresponding trait measure was somewhat lower than for the anxiousness self-rating. Nevertheless, a Steiger's (1980) test of correlation differences revealed that the bipolar angriness self-rating correlated marginally higher with the trait form of the State Trait Anger Expression Inventory, $r = .45$, than the bipolar anxiousness self-rating, $r = .23$, $t_{(97)} = 1.65$, $p = .05$ (one-tailed). Moreover, the trait form of the State Trait Anger Expression Inventory did not only correlate with the bipolar anxiousness self-rating but also with other direct anxiousness measures. This indicated a lack of discriminant validity for the trait form of the State Trait Anger Expression Inventory rather than for the bipolar anxiousness self-rating. This may further account for the only intermediate correlation between the bipolar angriness self-rating and the trait form of the State Trait Anger Expression Inventory. As a result, as expected by *Hypothesis 2*, convergent and

discriminant validity with established measures were shown for both bipolar self-ratings. This validated the word material I used as attributes within the IATs, at least at the level of direct measures.

5.4.3 Descriptive Statistics for the Anxiousness, the Angriness IAT, and the Behavioral Measures

Before I explore the correlations of the IATs and the behavioral measures, I will discuss briefly their descriptive statistics. The mean raw score (in milliseconds) of the *anxiousness IAT* was M = -171.1, SD = 156.9, and ranged from -640.6 to 179.2. Only 9 (6 female, 3 male) out of 100 participants had positive IAT scores. Thus, most of the participants were quicker to combine Me+self-confident and Others+anxious than for the reverse mapping. The mean raw score of the *angriness IAT* was M = -186.6, SD = 133.2, and ranged from –533.3 to 161.0. Only 4 (1 female, 3 male) out of 100 participants had positive scores. Thus, most of the participants were quicker to combine Me+self-controlled and Others+angry than for the reverse mapping. Mean error rates were for the anxiousness IAT M = 4.2%, SD = 2.6%, and for the angriness IAT M = 3.6%, SD = 2.3%. In any IAT, no participant had error rates higher than 15% or more than 10% of the latencies faster than 300 ms. The distributions of the improved and individually standardized D measures were not even marginally different from a normal distribution in both IATs, $Z < 1$. Internal consistency was computed across the two test halves and was acceptable for the anxiousness IAT, $\alpha = .72$, but somewhat unsatisfactory for the angriness IAT, $\alpha = .66$.

The descriptive statistics of the *behavioral measures* are depicted in Table 14. It should be noted that the reliability of several behavioral anger indicators was below .70. Nevertheless, the reliabilities of the global observer judgments for anxiety and anger were completely satisfactory.

Table 14

Descriptive Statistics of the Behavioral Measures in Study 2

Behavioral anxiousness measure (range of scores)	*N*[a]	*M*	*SD*	Range	Reliability[b]
Observer anxiety judgment (1-7)	100	3.27	1.06	1.33-6.42	.89
Anxious voice rating (1-7)	100	3.68	1.01	1.83-6.33	.72
Nervous mouth movements (per minute)	100	5.49	2.72	0.33-15.66	.87
Facial adaptor duration (%)	100	1.69	4.97	0 – 27	.99
Body adaptor duration (%)	100	13.52	23.10	0 – 96	.99
Illustrator duration (%)	100	8.55	14.66	0 – 72	.96
Behavioral angriness measure (range of scores)					
Observer anger judgment (1-7)	77	3.80	.83	1.75-6.08	.87
Angry voice rating (1-7)	77	2.96	.92	1.50-5.67	.69
Lips tight (per minute)	76	2.36	1.82	0-7.94	.65
Lips pressed (per minute)	76	0.44	.64	0-3.60	.82
Brows lower (per minute)	76	0.39	.69	0-2.77	.64

Note. *M*, *SD* and range refer to raw scores, reliabilities to log-transformed scores in the case of the behavior codings.

[a] sample size is smaller for anger indicators because participants, who realized that the anger induction was part of the experiment, had to be excluded from the analyses of the anger induction. One participant stood up during the anger induction, and his facial expression could, therefore, not be coded.

[b] agreement α of 4 observers for observer judgments, and of 2 observers for voice ratings, correlation *r* between 2 independent codings for behavior codings ($n = 20$).

5.4.4 Correlations of Direct, Indirect and Behavioral Measures with Social Desirability

The correlations between the direct anxiousness and angriness measures and social desirability are depicted in the last column of Table 13. As expected by *Hypothesis 3*, almost all direct measures showed small to intermediate correlations with social desirability. On the contrary, the anxiousness and the angriness IAT were not significantly correlated with social desirability, $r = .02$, $r = .-08$. Likewise, the observer anxiety and anger judgments showed no substantial correlations with social desirability, $r = .06$, $r = .-05$.

Unexpectedly, the bipolar anxiousness self-rating did not, in contrast to the Manifest Anxiety Scale and the trait form of the State Trait Anxiety Inventory, significantly correlate with social desirability, $r = -.08$. A possible explanation might be that, although the *anxious* pole of the bipolar self-rating represents socially *undesirable* traits, the opposed *self-confident* pole does not clearly stand for socially *desirable* traits. Conceptually, social desirability scales aim to assess the degree to which persons describe themselves in socially desirable terms (e.g., "I am always polite."). Therefore, social desirability is strongly related to agreeableness. Thus, although traits like self-confident, secure and unconcerned have a clear positive valence (cf. Chapter 0), these traits do not refer to socially adaptive and considerate behaviors. In contrast, the *angry* pole of the bipolar angriness self-rating clearly represents socially *undesirable* traits, and the opposed *self-controlled* pole clearly stands for socially *desirable* traits. This was consequently reflected in the negative correlation between the bipolar angriness self-rating and social desirability, $r = -.30$, $p < .01$.

5.4.5 Moderation of the Relationship between Direct and Indirect Measures by Social Desirability

To examine whether social desirability moderated the relationship between direct and indirect measures according to *Hypothesis 4* I conducted stepwise multiple regression analyses. In these regressions, the direct anxiousness or angriness measures were the criteria. In the first step, social desirability and the anxiousness or angriness IAT were entered as predictors. In the second step, the interaction term (cross product) of both variables (each scored as deviation of the original scale from its own mean) was added as a predictor. The results for every direct anxiousness and angriness measure are depicted in

Table 15. As indicated by the zero-order correlations (Table 13), direct anxiousness measures were predicted by the anxiousness IAT, and direct anxiousness and angriness measures were predicted by social desirability in almost every case. However, when the interaction term of social desirability and the IAT was entered in step 2, there was never a significant increment in the explained variance. Thus, as expected from Hypothesis 4 social desirability did not moderate the relationship between either direct and indirect anxiousness or direct and indirect angriness measures. The results were the same when regression analyses were conducted separately for both groups of different IAT order.

Table 15

Moderation of the Relationship between Direct and Indirect Measures by Social Desirability

	Step 1		Step2				
			IAT[b]	SD			IAT[b] × SD
Direct measure	R^2	$F_{(2, 97)}$	β	β	ΔR^2	$F_{(1, 96)}$	β
Bipolar anxiousness self-rating	.071*	3.72*	.26*	-.08	.000	.04	-.02
Speaking Anxiety Emotionality	.025	1.23	-.01	-.16	.000	.00	.00
Speaking Anxiety Worries	.086*	4.59*	.18^{+}	-.24*	.019	2.00	.14
State Trait Anxiety Inventory[a]	.094**	5.03**	.18^{+}	-.25*	.005	.51	.07
Manifest Anxiety Scale	.135***	7.60***	.22*	-.30**	.000	.00	.00
Bipolar angriness self-rating	.097**	5.22**	.09	-.29**	.000	.05	-.02
State Trait Anger Expression Inventory[a]	.118**	6.51**	.00	-.34***	.000	.01	.01

Note. N = 100. IAT = Implicit Association Test, SD = social desirability. [a] trait form. [b] Anxiousness IAT for prediction of direct anxiousness measures and angriness IAT for prediction of direct angriness measures. $^{+}p < .10$ $*p < .05$ $**p < .01$ $***p < .001$.

5.4.6 Zero Correlation between Anxiousness and Angriness

Conceptualizing anxiousness and angriness as orthogonal dimensions, it was expected by *Hypothesis 5* that both these traits were not correlated. Hypothesis 5 was confirmed for the correlation between the bipolar anxiousness and angriness self-rating, $r = -.08$, *n.s.*, and the observer anxiety and anger judgment, $r = .00$, *n.s.*. Nevertheless, the trait form of the State Trait *Anger* Expression Inventory showed intermediate correlations with the trait form of the State Trait *Anxiety* Inventory and the Manifest *Anxiety* Scale, and small correlations with the bipolar *anxiousness* self-rating (see Table 13). This replicated the results of some previous studies (Schwenkmezger et al., 1992), that showed that the State Trait Anger Expression Inventory was correlated with anxiousness because individuals high in neuroticism are more concerned with their anger expression than those individuals who are emotionally stable. When anxiousness and angriness were conceptualized as orthogonal dimensions, the bipolar self-ratings did not correlate with each other, and the angriness self-rating was not correlated with any direct anxiousness measure.

In contrast, *Hypothesis 5* was not confirmed for the correlation between the anxiousness and the angriness IAT that was significantly positive, $r = .32$ $p < .01$. Moreover, order effects affected this correlation. The sequence of the anxiousness and the angriness IAT was counterbalanced across participants such that two groups with different IAT orders could be compared with each other. In the group that completed the anxiousness IAT as first test, both IATs were substantially correlated, $r = .49$, $p < .001$, whereas they were not even marginally correlated in the group that completed the angriness IAT first, $r = .17$, *n.s.*. This correlation difference was marginally significant, $z = 1.77$, $p < .10$ (two-tailed). The discrepancy might not be attributed to sample effects, as anxiousness and angriness were neither correlated for the bipolar self-ratings nor for the observer judgments in both groups, all $r < .17$, *n.s.*.

A possible explanation might be that anxiousness normally shows higher correlations with neuroticism than angriness. This was also the case in the pilot study that was performed to select the bipolar items. In that pilot study ($N = 42$), anxiousness and neuroticism were strongly correlated, $r = .82$, $p < .001$, whereas angriness and neuroticism showed only a weak correlation, $r = .21$, $p = .18$. Working on the anxiousness IAT, participants could have possibly developed a classification heuristic, discriminating anxious versus self-confident as neurotic versus non-neurotic or even as positive versus negative attributes. In other words, participants recoded the IAT task because a discrimination of positive versus negative is easier than a discrimination of anxious versus self-confident (cf. De Houwer 2003a). This *task-recoding* was salient during the anxiousness IAT. Upon completion of the anxiousness IAT the task-recoding could have been transferred onto the angriness IAT, which would lead to a positive correlation between both IATs. In contrast, the angriness IAT is less likely to elicit to a positive-negative task-recoding, because angry versus self-controlled is less associated with neuroticism. Consequently, when the angriness IAT was the first test, the participants did not use a positive-negative classification, and the IATs did not correlate with each other.

To examine whether a *positive-negative dimension* is more salient in anxious versus self-confident than in angry versus self-controlled judgments, participants of two different groups rated the valence of the IAT stimuli. Instructions for the self-relevant group (41 undergraduate psychology students) asked to estimate how positive or negative one would rate a trait if it was one's own. This was done because the self-concept IATs ask participants to combine 'Me' with personality traits, for example, anxious. "Anxious" may be judged more negatively when it refers to oneself rather than to anxiousness in general. Instructions for the control group (10 PhD psychology students) simply asked respondents to rate the positiveness or negativeness of traits in general. In both groups, the anxiousness and angriness traits were presented in a paper-pencil questionnaire, and respondents judged the valence of those traits on a 7-point scale (negative [---] [--] [-] [0] [+] [++] [+++] positive). Answers were coded such that higher values indicated more positive valence. The results are shown in Table 16.

Table 16

Valence Ratings of the IAT Stimuli from Two Different Samples

	Undergraduates (n = 41)			PhD students (n = 10)			Group difference (df = 49)		
Attributes	*M*	*SD*	Range	*M*	*SD*	Range	*t*	*p*	*d*
Anxious (ängstlich)	2.68	1.15	1-6	2.20	.79	1-3	1.25	.22	.36
Timid (furchtsam)	2.49	1.08	1-5	2.20	.92	1-4	.78	.44	.22
Insecure (unsicher)	2.20	.84	1-4	2.50	1.18	1-5	-.95	.35	-.27
Worried (besorgt)	3.78	1.44	1-6	3.80	1.55	3-8	-.04	.97	-.01
Overly cautious (übervorsichtig)	2.17	1.00	1-5	2.20	.63	1-3	-.09	.93	-.03
Mean anxious attributes	2.66	.83	1.2-5.0	2.58	.60	1.6-3.6	.30	.77	.09
Self-confident (sicher)	6.02	.82	4-7	6.10	.74	5-7	-.27	.79	-.08
Daring (wagemutig)	4.85	1.20	3-7	4.10	1.10	3-6	1.81	.08	.52
Secure (selbstvertrauend)	6.37	.66	5-7	6.60	.52	6-7	-1.04	.30	-.30
Unconcerned (sorglos)	4.24	1.56	1-7	4.00	1.25	2-6	.46	.65	.13
Carefree (unbeschwert)	5.44	1.23	2-7	5.40	1.07	4-7	.09	.93	.03
Mean self-confident attributes	5.39	.69	3.6-6.6	5.24	.52	4.6-6.0	.62	.54	.18
Angry (ärgerlich)	3.07	1.27	1-6	3.00	1.33	1-6	.16	.87	.05
Hot-tempered (aufbrausend)	2.34	1.28	1-6	2.60	1.43	1-5	-.56	.58	-.16
Undercontrolled (unbeherrscht)	1.88	.87	1-4	1.90	.88	1-3	-.07	.94	-.02
Hot-headed (hitzköpfig)	2.83	1.30	1-7	2.20	1.03	1-4	1.42	.16	.41
Irritable (motzig)	1.95	1.09	1-6	1.80	.79	1-3	.41	.68	.12
Mean angry attributes	2.41	.70	1.4-4.6	2.30	.60	1.6-3.4	.47	.64	.14
Self-controlled (kontrolliert)	4.76	1.37	2-7	4.30	1.34	2-6	.95	.35	.27
Thoughtful (bedächtig)	4.73	.92	3-7	4.90	.88	3-6	-.52	.60	-.15
Self-disciplined (selbstbeherrscht)	4.98	1.19	3-7	5.00	1.15	2-6	-.06	.95	-.02
Adaptable (fügsam)	2.46	1.16	1-6	2.40	.52	2-3	.17	.87	.05
Calm (friedlich)	5.54	1.05	3-7	5.70	1.16	3-7	-.43	.67	-.12
Mean self-controlled attributes	4.49	.61	3.4-6.0	4.46	.65	3.4-5.4	.15	.88	.04

Note. The scale format was a 7-point scale with 1 indicating negative, 4 indicating neutral, and 7 indicating positive valence.

As it can be seen from Table 16, the valence of the traits was not judged differently by the undergraduates and the PhD students, although the undergraduates rated the valence as if the traits were their own. “Daring” was judged marginally more positive by the undergraduates. However, given the amount of tests performed, this might may very well be a chance finding. It should be noted that the design of this valence check confounded group membership (undergraduates versus PhD students) and instruction (self-relevant versus control). However, it is unlikely to expect that the PhD students and the undergraduates had a different self-concept in anxiousness or angriness. Thus, the direct valence estimates seemed to be unaffected by the instruction to judge the traits as if one’s own. More importantly, the group differences did not consistently point in the same direction, neither for the positively ($M > 4$) nor for the negatively ($M < 4$) evaluated traits. Since the sample size of the undergraduate group was larger, and the undergraduate group is more similar to the sample of Study 2, the results of this group will be discussed in regards to the valence estimates.

As one may recall, the category label of the attribute concept was anxious versus self-confident for the anxiousness IAT, and angry versus self-controlled for the angriness IAT, respectively. Since the category label has a chief influence on the IAT effect (cf. Chapter 2.4.2), the valence estimates for the category labels as well as for the category means were compared. Concerning the labels, anxious was rated more negatively than self-confident, $d = 3.17$. (The effect size d for repeated measures was computed as $\sqrt{2}(M_1 - M_2)/SD$ where SD is the standard deviation of the difference scores; see Cohen, 1988). Similarly, angry was rated more negatively than self-controlled, $d = 1.33$, but the effect size was less than half than for anxious versus self-confident. Thus, a positive-negative dimension was stronger for anxious versus self-confident than for angry versus self-controlled. More importantly, self-confident was also rated more positively than self-controlled, $d = 1.17$. “Anxious” was not judged more negatively than angry, $t_{(40)}$ 1.41, $p = .17$, $d = .31$, although the effect pointed in the expected direction (see the first column of Table 16).

Concerning the category means, the five anxious attributes were rated more negatively than the five self-confident attributes, $d = 3.03$. However, the five angry attributes were also rated more negatively than the five self-controlled attributes, $d = 3.31$. Thus, at the level of category means, a positive-negative dimension was as strong for anxious versus self-confident as for angry versus self-controlled. The five self-confident

attributes were judged more positively than the five self-control attributes, $d = 1.50$. The five anxious attributes were not judged more negatively than the five angry attributes, $t_{(40)} = -1.35$, $p = .18$, $d = -.31$, and the effect did not even point in the expected direction (see the first column of Table 16).

In summary, the positive-negative difference was stronger for anxious versus self-confident than for angry versus self-controlled. This was true only at the level of the category labels but not at the level of the category means. Nevertheless, self-confident was judged more positively than self-controlled concerning the category labels as well as the category means. More importantly, within the self-control attributes, one attribute (adaptable) was judged negatively when it was tested against the neutral scale midpoint, $t_{(40)} = -8.45$, $p < .001$. In contrast, none of the anxious attributes was judged positively, and none of the self-confident attributes was judged negatively.

Altogether, a positive-negative dimension was less clear in the angriness IAT than in the anxiousness IAT concerning both the level of category labels and the level of category exemplars. Consequently, a positive-negative dimension was less salient within the angriness IAT. As a result, there might have been a transfer effect from the anxiousness IAT on the angriness IAT, but not vice versa. In the next section, I present the effects of different IAT order on the correlations of the anxiousness and the angriness IAT concerning the correlations with direct self-ratings and the observer judgments.

5.4.7 Order Effects on IAT Correlations

It was expected by *Hypothesis 6* that the second IAT tended to be less valid than the first IAT. Given the transfer effect from the anxiousness on the angriness IAT this should be especially true for the angriness IAT. Table 17 depicts the overall correlations and the correlations by IAT order for both IATs. Concerning the anxiousness IAT, all correlations with direct anxiousness measures declined in the second test, except for the trait form of the State Trait Anxiety Inventory. When the anxiousness IAT was the first test, it showed significant or marginally significant correlations with several direct anxiousness measures, whereas these correlations were not even marginally significant when it was the second test. This correlation decrease was marginally significant for the subscale Worries of the Speaking Anxiety Questionnaire, $z = 1.39$, $p < .10$ (one-tailed), and not even marginally significant for the correlations of the anxiousness IAT with other direct anxiousness measures, all $z < 1.12$, *n.s.* (one-tailed). Contrary to Hypothesis 6, the correlation with the

observer anxiety judgment tended to be higher, and was significant only when the anxiousness IAT was the second test. However, this correlation difference was small and nonsignificant, $z = -.58$, *n.s.* (one-tailed). In regards to the anxiousness IAT, a pattern of reduced validity for the second test was confirmed for the correlations with direct anxiousness measures but not for the correlation with the observer judgment.

Table 17

Overall and Correlations by IAT Order for the Anxiousness and the Angriness IAT (Study 2)

	Anxiousness IAT			Angriness IAT		
	Overall	1st test	2nd test	Overall	1st test	2nd test
Angriness IAT	.32**	.49***	.17	-	-	-
Bipolar anxiousness self-rating	.25*	.28*	.23	-.04	-.18	.15
Bipolar angriness self-rating	-.03	-.06	.03	.11	.16	.06
Speaking Anxiety Emotionality	-.01	.02	-.08	-.03	-.24^{+}	.20
Speaking Anxiety Worries	.17^{+}	.27^{+}	-.01	.05	-.16	.27^{+}
Manifest Anxiety Scale	.21*	.31*	.09	.00	-.15	.22
State Trait Anxiety Inventory[a]	.17^{+}	.17	.20	.02	-.06	.12
State Trait Anger Expression Inventory[a]	-.01	-.01	.01	.03	-.16	.25^{+}
Social Desirability	.02	-.07	.16	-.08	.02	-.20
Observer anxiety judgment	.26**	.22	.33*	-.07	-.05	-.09
Observer anger judgment	-.09[b]	-.04[c]	-.17[d]	-.11[b]	-.23[d]	.00[c]

Note. $N = 100$ for overall correlations, $n = 50$ for correlations by different IAT order. IAT = Implicit Association Test. [a] trait form. [b] $n = 77$. [c] $n = 41$. [d] $n = 36$.
$^{+}p < .10$ $*p < .05$ $***p < .001$.

In regards to the correlations of the angriness IAT, the pattern was less clear. This might be due to the fact that the angriness IAT showed only small convergent validity with direct angriness measures already in the first test. As expected by Hypothesis 6, the correlation with the bipolar angriness self-rating decreased when the angriness IAT was the second test. However, this correlation difference was small and nonsignificant, $z = .49$, *n.s.* (one-tailed). Moreover, the opposite was true for the trait form of State Trait Anger

Expression Inventory. When the angriness IAT was the second test, its correlation with the trait form of the State Trait Anger Expression Inventory was marginally significant and higher than in the first test, $z = -2.02$, $p < .05$ (two-tailed). The correlation with the observer anger judgment was even nonsignificantly negative when the angriness IAT was the first test and zero when it was the second test, but did not differ significantly between both groups, $z = -1.14$, *n.s.* (one-tailed). Thus, a pattern of reduced convergent validity for the second test was not found for the angriness IAT, and validity was small in both cases.

With regard to discriminant validity, the anxiousness IAT did not correlate with direct angriness measures in any case. In contrast, the angriness IAT correlated with the anxiousness IAT *and* tended to correlate with direct anxiousness measures, when it was the second test. Thus, there was a marginally positive correlation with the subscale Worries of the Speaking Anxiety Scale (see Table 17). On the other hand, these correlations tended to be negative, when the angriness was the first test. Specifically, there was a marginally negative correlation with the subscale Emotionality of the Speaking Anxiety Scale. Possibly, for some participants, the category self-controlled, that was the opposite of angry within the angriness IAT, was more related with high rather than low anxiousness. Therefore, the small negative correlation between the angriness IAT and direct anxiousness measures might have appeared. In contrast, when the angriness was the second test, it tended to positively correlate with direct anxiousness measures due to the assumed transfer effect from the anxiousness IAT onto the angriness IAT. The correlation differences (two-tailed tests) between both groups were significant for the two subscales of the Speaking Anxiety Scale (in both cases $z > 2.12$, $p < .05$), marginally significant for the Manifest Anxiety Scale ($z = 1.82$, $p < .10$), and nonsignificant for the bipolar anxiousness self-rating and the trait form of the State Trait Anxiety Inventory (in both cases $z < 1.61$, *n.s*). Consequently, the increase in correlation with direct anxiousness measures provides further evidence for a transfer effect from the anxiousness IAT on the angriness IAT.

The positive correlations of direct anxiousness measures with the angriness IAT in the second test might also lead to the positive correlation between the angriness IAT and the trait form of the State Trait Anger Expression Inventory. Table 13 shows that the trait form of the State Trait Anger Expression Inventory was positively correlated with direct anxiousness measures. Thus, the positive correlation between the angriness IAT and the trait form of the State Trait Anger Expression Inventory might be mediated by the correlation of both measures with direct anxiousness measures. However, when the

correlation between both measures was controlled for their correlation with direct anxiousness measures the partial correlation was only a little smaller than the zero-order correlation, $r = .20$, *n.s.* versus $r = .25$, $p < .10$. Thus, the angriness IAT seemed to show at least some convergent validity with the trait form of the State Trait Anger Expression Inventory.

Altogether, the anxiousness IAT showed a pattern of reduced validity for the second test with respect to direct measures but not for the observer anxiety judgment. The angriness IAT showed small convergent validity in general, and was affected by a transfer effect from the anxiousness IAT. This led to positive correlations between both IATs, and a trend to positive correlations between the angriness IAT and direct anxiousness measures.

5.4.8 Prediction of the State and the Behavioral Measures by Direct and Indirect Measures

In this section, I report the results of hierarchical regression analyses that explored whether state and behavioral measures of anxiety and anger were predicted by direct and indirect measures. According to *Hypothesis 7*, it was expected that self-reported state measures were predicted by self-reported trait measures, and that the IATs added incremental validity to self-reported measures to the prediction of behavior.

To examine the prediction of *anxiety* I performed separate hierarchical regressions with self-reported state anxiety and behavioral anxiety as criteria. Predictors were direct and indirect anxiousness measures, as well as direct state anxiety and its change when behavioral anxiety was the criterion. Direct measures (the bipolar anxiousness self-rating, the subscales Emotionality and Worries of the Speaking Anxiety Questionnaire, the trait form of the State Trait Anxiety Inventory, the Manifest Anxiety Scale, plus, for the prediction of anxious behavior, the bipolar state anxiety self-rating and its change) were entered in one step, and the anxiousness IAT was entered in the other step.

Table 18

Predictions of the State Anxiety Measures and the Behavioral Anxiety Indicators by Direct Measures and the Anxiousness IAT

	Hierarchical regression	
	Step 1: Direct measures[a]	Step 2: Anxiousness IAT
Measure	R^2	ΔR^2
Bipolar state anxiety self-rating		
Speech	.448***	.000
Change (speech minus baseline)	.131*	.001
Behavioral anxiety indicators		
Observer anxiety judgment	.171*	.072**
Anxious voice rating	.131^{+}	.043*
Facial adaptor duration	.083	.003
Body adaptor duration	.080	.006
Illustrator duration	.055	.004
Nervous mouth movements (frequency)	.054	.000

Note. N = 100. IAT = Implicit Association Test.

[a] For the regression analysis on direct state anxiety all direct anxiousness measures (the bipolar anxiousness self-rating, the subscales Emotionality and Worries of the Speaking Anxiety Questionnaire, the trait form of the State Trait Anxiety Inventory, and the Manifest Anxiety Scale) were entered. For regression analysis on behavioral anxiousness indicators all direct anxiousness plus the state anxiety measures (bipolar self-rating and its change) were entered. $^{+}p < .10$ $*p < .05$ $**p < .01$ $***p < .001$.

In order to evaluate the contribution of every single predictor and to control for suppressor effects, I carried out different regressions considering the following points: (a) The contribution of each direct measure was individually analyzed in a separate regression entering the direct measure in one step, and the IAT in the other step. (b) Both orders of these hierarchical regressions were organized such that the direct measure was entered in Step 1 and the IAT in Step 2, as well as the opposite order of both steps . (c) Predictive validity of the IAT was inspected separately for both groups of different IAT order

(anxiousness IAT as first versus as second test). To avoid accumulation of α-error I first performed overall hierarchical regressions entering all direct measures in Step 1 and the IAT in Step 2. Then, I performed further analyses following points (a) to (c). To report the results for each criterion I begin with the overall analysis, as depicted in Table 18. Results are then outlined more clearly with the findings of points (a) to (c). I conclude with examining the standardized βs of all predictors in Step 2 of the overall analysis. To keep these analyses manageable I did not consider marginally significant results.

As it can be seen in the first row of Table 18, direct anxiousness measures significantly accounted for self-reported *state anxiety* immediately before the speech, whereas the anxiousness IAT did not. This was (a) true for every single direct anxiousness measure, (b) independent of the regression order, and (c) not affected by different IAT orders. However, although all direct anxiousness measures share significant portions of variance with self-reported state anxiety (see Table 19), they did not independently contribute to the criterion. When all predictors were entered into the overall regression (Step 2 in Table 18), only the bipolar anxiousness self-rating and the Emotionality subscale of the Speaking Anxiety Scale were significant predictors, $\beta = .41$, $t = 3.42$, $p < .001$, $\beta = .43$, $t = 3.89$, $p < .001$, all others $|\beta| < .09$, $|t| < .70$, *n.s.*.

Direct anxiousness measures also predicted *state anxiety change* after the anxiety induction, whereas the anxiousness IAT did not. This was (a) only true for the Emotionality subscale of the Speaking Anxiety Questionnaire, and independent of (b) regression order and (c) IAT order. Accordingly, only the Emotionality subscale accounted for the increase in state anxiety in the overall analysis (Step 2 in Table 18), $\beta = .46$, $t = 3.22$, $p < .01$, all others $|\beta| < .27$, $|t| < 1.81$, *n.s.*.

Table 19

Correlations of State Anxiety (Study 2)

	Bipolar state anxiety self-rating	
State anxiety	Speech	Change (speech minus baseline)
Bipolar state anxiety self-rating (speech)	-	.66***
Anxiousness measures		
Bipolar anxiousness self-rating	.56***	.09
Speaking Anxiety Emotionality	.52***	.29**
Speaking Anxiety Worries	.40***	.09
Manifest Anxiety Scale	.47***	.13
Trait form of the State Trait Anxiety Inventory	.45***	.03
Anxiousness IAT	.11	.00
Behavioral anxiety		
Observer anxiety judgment	.38***	.23*
Anxious voice rating	.30***	.14
Facial adaptor duration	.07	-.02
Body adaptor duration	.16	.07
Illustrator duration	.09	.19^{+}
Nervous mouth movements (frequency)	.15	.06

Note. $N = 100$. IAT = Implicit Association Test.

$^{+}$ $p < .10$ * $p < .05$ ** $p < .01$ *** $p < .001$.

As it can be seen in Table 18 the *observer anxiety judgment* was predicted by direct and indirect measures. (a) This was true for the bipolar anxiousness self-rating, the Emotionality subscale, the bipolar state anxiety self-rating, and the change in state anxiety. The trait form of the State Trait Anxiety Inventory and the Manifest Anxiety Scale contributed marginally to the observer anxiety judgment. The Worries subscale was not even a marginal predictor. The anxiousness IAT accounted for the observer anxiety judgment independently from all direct measures. (b) When the anxiousness IAT was entered first, only the Emotionality subscale, the self-reported state anxiety and its change additionally contributed to the observer judgment. (c) As indicated by the zero-order

correlations (Table 20), the anxiousness IAT showed significant correlations with, and was, therefore, a significant predictor for the observer anxiety judgment when it was preceded by the angriness IAT. When the anxiousness IAT was the first test it marginally predicted the observer anxiety judgment. To conclude, only the Emotionality subscale, the bipolar state anxiety self-rating, and the anxiousness IAT were significant predictors in the overall analysis, $\beta = .33$, $t = 2.26$, $p < .05$, $\beta = .36$, $t = 2.05$, $p < .05$, $\beta = .29$, $t = 2.94$, $p < .01$, all others $|\beta| < .27$, $|t| < 1.81$, *n.s.*.

Table 20

Correlations of Behavioral Anxiety Measures in Study 2

		Anxiousness IAT			Explicit anxiousness			Speaking Anxiety	
Behavioral measure	Observer judgment	1st test	2nd test	Both	Bipolar self-rating	MAS	STAI	Emotio-nality	Worries
Observer judgment	-	.22	.33*	.26**	.22*	.19+	.19+	.29**	.15
Anxious voice rating	.61***	.23	.19	.22*	.14	.06	.13	.23*	.22*
Facial adaptor duration	-.14	-.11	-.01	-.07	-.01	-.16	-.09	-.02	-.07
Body adaptor duration	.26**	.02	.13	.06	.07	-.06	-.03	.07	-.05
Illustrator duration	.02	.23	-.15	.05	.00	-.01	-.05	-.05	-.11
Nervous mouth movements (frequency)	.22*	.00	.11	.04	.13	. 07	.15	-.03	.00

**Note*. N = 100 (n = 50 for different IAT orders). IAT = Implicit Association Test, MAS = Manifest Anxiety Scale, STAI = trait form of the State Trait Anxiety Inventory. * $p < .05$ ** $p < .01$ *** $p < .001$.

The *anxious voice* rating was marginally predicted by direct measures and significantly predicted by the anxiousness IAT. (a) With regards to single direct measures, the Worries and the Emotionality subscale, as well as the bipolar state anxiety self-rating were significant predictors. The anxiousness IAT marginally contributed to the anxious voice rating, when entered after the bipolar anxiousness self-ratings, the Worries subscale, or the state anxiety self-rating. The anxiousness IAT was always a significant predictor, when entered after any other direct measure. (b) When the direct measure was entered after the anxiousness IAT, the Emotionality subscale as well as the self-reported state anxiety were significant predictors and the Worries subscale was a marginal predictor. (c) As it

may be seen in Table 20, the anxiousness IAT significantly correlated with the anxious voice rating only when both groups with different IAT order were pooled. When the groups were inspected separately, sample sizes were smaller, and the small positive correlations failed to reach the significance criterion although the effect sizes were almost the same. This was also true for the regression analysis. Thus, the IAT was a significant predictor only when the groups with different IAT order were analyzed simultaneously. Finally, in the overall regression only the anxiousness IAT was a significant predictor $\beta = .22$, $t = 2.17$, $p < .05$. The bipolar state anxiety self-rating marginally accounted for the anxious voice rating, $\beta = .37$, $t = 1.98$, $p < .10$, all others $|\beta| < .27$, $|t| < 1.54$, *n.s.*.

The duration of *facial adaptors*, *body adaptors*, and *illustrators* as well as the frequency of *nervous mouth movements* was neither predicted by direct measures nor by the anxiousness IAT. (a) When entered as single predictors, only the increase in state anxiety was a marginal but surprisingly positive predictor for illustrator duration. However, the effect was only small and might be due to chance. All other direct measures were not even marginal predictors for any of the anxiety codings. This pattern was not affected by (b) regression or (c) IAT order. In the overall analyses, neither any direct measure nor the anxiousness IAT was a significant predictor, all $|\beta| < .19$, $|t| < 1.19$, *n.s.*. This was the case even though nervous mouth movements and the duration of body adaptors showed small correlations with the observer anxiety judgment (see Table 20). Thus, although the observers interpreted nervous mouth movements and body adaptors as anxious behavior, these codings were not related to self-reported anxiousness and anxiety measures or the anxiousness IAT (see Table 20). It should be noted that the observer anxiety judgments showed a large correlation with the anxiety rating of the participants' voices (see Table 20). Therefore, important anxiety indicators might not be found in the gestures or the facial expressions, but in the verbal expression and the sound of the participants' voices.

To summarize these findings, self-reported state anxiety was predicted by direct anxiousness measures but not by the anxiousness IAT. This confirmed *Hypothesis 7*. Important predictors for state anxiety were the bipolar anxiousness self-rating and the Emotionality subscale of the Speaking Anxiety Scale. In contrast, the anxiousness IAT added incremental validity over direct measures to the prediction of the observer anxiety judgment and the anxious voice rating. This confirmed, again, *Hypothesis 7*. It should be noted that the observer anxiety judgment and the anxious voice rating was also predicted

by direct measures. Important predictors were, again, the Emotionality subscale and the self-reported state anxiety. Codings of anxious behavior were neither predicted by direct measures nor by the anxiousness IAT.

To examine the prediction of *anger*, I carried out the same hierarchical regressions as for the prediction of anxiety but with self-reported state anger and behavioral anger as criteria. Predictors were direct and indirect angriness measures as well as direct state anger and its change when behavioral anger was used as the criterion. Again, direct measures (the bipolar angriness self-rating, the trait form of the State Trait Anger Expression Inventory, plus, for the prediction of angry behavior, the bipolar state anger self-rating and its change) were entered in one step, while the angriness IAT was entered in the other step. Results are reported considering the same aspects as for the prediction of anxiety. I start with the results of the overall analysis as depicted in Table 21. Then, I explore (a) the contribution of single direct measures, (b) different regression orders, and (c) different IAT orders. Finally, I examine the standardized βs of all predictors in Step 2 of the overall analysis.

As it is shown in Table 21, the bipolar *state anger* self-rating after the computer crash and the *state anger change* as compared to the baseline were neither predicted by direct angriness measures nor by the angriness IAT. This was (a) the same for the trait form of the State Trait Anger Expression Inventory and the bipolar angriness self-rating, and not affected by (b) regression or (c) IAT order. Thus, in both overall analyses neither direct angriness measures nor the angriness IAT were significant predictors, all $|\beta| < .18$, $|t| < 1.64$, *n.s.*.

Table 21

Predictions of the State Anger Measures and the Behavioral Anger Indicators by Direct Measures and the Angriness IAT

	Hierarchical regression	
	Step 1: Direct measures	Step 2: Angriness IAT
Measure	R^2	ΔR^2
Bipolar state anger self rating		
Computer crash	.044	.000
Change (computer crash minus baseline)	.003	.006
Behavioral angriness indicators		
Observer anger judgment	.224***	.015
Angry voice rating	.154*	.014
Lips tight (frequency)	.029	.144***
Lips pressed (frequency)	.020	.023
Brows lower (frequency)	.110^{+}	.003

Note. N = 77. IAT = Implicit Association Test.

[a] For the regression analysis on direct state anger all direct angriness measures (the bipolar angriness self-rating and the trait form of the State Trait Anger Expression Inventory) were entered. For regression analysis on behavioral angriness indicators all direct angriness plus the state anger measures (bipolar self-rating and its change) were entered. $^{+}p < .10$ $^{*}p < .05$ $^{***}p < .001$.

The *observer anger judgment* was predicted by direct measures but not by the angriness IAT. (a) This was true for the trait form of the State Trait Anger Expression Inventory and for the bipolar angriness self-rating but not for self-reported state anger or its change. This pattern was not affected by (b) regression or (c) IAT order. The only significant predictor in the overall analysis was the bipolar angriness self-rating, $\beta = .34$, $t = 2.94$, $p < .01$. The trait form of the State Trait Anger Expression Inventory and the self-reported change in state anger were only marginally significant predictors, $\beta = .21$, $t = 1.87$, $p < .10$, $\beta = .25$, $t = 1.90$, $p < .10$. The self reported state anger did not significantly account for the observer anger judgment, $\beta = -.12$, $t = -.87$, *n.s.*.

Table 22

Correlations of State Anger (Study 2)

	Bipolar state anger self-rating	
State anger	Computer crash	Change (computer crash minus baseline)
Bipolar state anger self-rating (computer crash)	-	.60***
Angriness Measures		
Bipolar angriness self-rating	.18	-.04
State Trait Anger Expression Inventory[a]	.17	.02
Angriness IAT	-.06	-.08
Behavioral Anger		
Observer anger judgment	.14	.18
Angry voice rating	.28*	.28*
Lips tight (frequency)	.03	-.03
Lips pressed (frequency)	-.07	-.08
Brows lower (frequency)	-.04	.20^{+}

**Note*. N = 100, n = 77 for behavioral anger measures. IAT = Implicit Association Test. [a] Trait form. $^{+}$ $p < .10$ * $p < .05$ *** $p < .001$.

Similarly the *angry voice rating* was predicted by direct measures but not by the angriness IAT. (a) When direct measures were analyzed individually, the bipolar angriness self-rating, the bipolar state anger self-rating, and its change significantly accounted for the angry voice rating. In contrast, the trait form of the State Trait Anger Expression Inventory failed to be a significant predictor. This pattern was not affected by (b) regression or (c) IAT order. In the overall analysis the bipolar angriness self-rating was the only even so marginally significant predictor, $\beta = .24$, $t = 1.96$, $p < .10$, whereas all others were $|\beta| < .23$, $|t| < 1.66$, *n.s.*.

The frequency of putting the *lips tight* was predicted by the angriness IAT but not by direct measures. However, contrary to expectations, the angriness IAT and the frequency of tight lips were negatively correlated (see Table 23). When (a) direct measures were inspected individually or (b) the regression order was varied, results were the same. Nevertheless, this pattern was true only when (c) the angriness IAT was the first test. As already indicated by the zero-order correlations (Table 23), the frequency of tight lips and

the angriness IAT share common portions of variance only when the angriness IAT was completed before the anxiousness IAT. Concerning the overall analysis, only the angriness IAT was a significant, although negative predictor, $\beta = -.38$, $t = -3.48$, $p < .001$, all others were $|\beta| < .14$, $|t| < 1.19$, *n.s.*.

Table 23

Correlations of Behavioral Anger Measures in Study 2

Behavioral measure	Observer judgment	Angriness IAT			Explicit angriness	
		1st test	2nd test	Both	Bipolar self-rating	STAXI
Observer anger judgment		-.23	.00	-.05	.38**	.33**
Angry voice rating	.51***	-.17	-.10	-.10	.25*	.16
Lips tight (frequency)	.08	-.55**	-.21	-.34**	.05	.16
Lips pressed (frequency)	-.05	-.28+	.00	-.13	.00	-.11
Brows lower (frequency)	.08	-.12	.00	-.12	-.07	-.21+

Note. $n = 77$ ($n = 36$ for angriness IAT as first test, and $n = 41$ for angriness IAT as second test). IAT = Implicit Association Test. $^{+}p < .10$ $^{*}p < .05$ $^{**}p < .01$ $^{***}p < .001$.

The frequency of pressing the lips together (*lips pressed*) was neither predicted by direct measures nor the angriness IAT. This pattern (a) was true for every single direct measure. When the angriness IAT (b) was entered before the direct measures and (c) was the first test, it marginally accounted for the frequency of pressed lips. However, as it is indicated by the correlations in Table 23, the angriness IAT was then, once more, a negative predictor. In the overall analysis, none of the predictors was even marginally significant, all $|\beta| < .15$, $|t| < -1.31$, *n.s.*.

The frequency of frowns (*brows lower*) was marginally predicted by direct measures but not by the angriness IAT. (a) Concerning single direct measures, this was true for the change in bipolar state anger and the trait form of the State Trait Anger Expression Inventory. However, the prediction was positive only for the former and again surprisingly negative for the latter (see the zero-order in Table 22 and in Table 23). This pattern was neither affected by (b) regression nor (c) IAT order. In the overall regression, only the change in state anger was a significant predictor, $\beta = .32$, $t = 2.27$, $p < .05$, all

others $|\beta| < .20$, $|t| < 1.64$, *n.s.*. Since the frequency of frowns did not correlate with the observer anger judgment, and the effects of the direct measures were small and contradictory, these results might be due to chance.

To summarize these findings, *Hypothesis 7* was not confirmed with regard to the prediction of state anger through direct angriness measures. Hypothesis 7 was also not confirmed with regard to the incremental validity of the angriness IAT for the prediction of angry behavior. The negative correlation of the angriness IAT with the frequencies of tight and pressed lips was contrary to expectations, and true only when the angriness IAT was the first test. Moreover, as it can be seen from Table 23, none of the anger codings were correlated with the observer anger judgment. Therefore, the anger codings might not be valid indicators for angry behavior. However, the observer anger judgment correlated substantially with the anger rating of the participants' voices (see Table 23). Thus, important anger indicators might not be found in the facial expressions, but in the verbal expression and the sound of the participants' voices. Finally, the observer anger judgment and the angry voice rating were only predicted by direct measures, whereby the bipolar angriness self-rating was the most important predictor.

5.5 Discussion

In the Discussion section, I first summarize the main findings of Study 2, and then briefly refer to gender differences. Subsequently, I discuss the differences between direct and indirect measures concerning the prediction of anxious and angry behavior. Finally, I refer to the conceptualization of angriness within the present study.

5.5.1 Summary of the main findings

Study 2 explored the psychometric properties of an anxiousness and an angriness IAT. Thereby, the sequence of the IATs was counterbalanced. The IATs' convergent and discriminant validity was examined both for self-reported anxiousness and angriness, as well as for anxious versus angry behavior after emotion inductions. Study 2 tested seven hypotheses.

First, the efficacy of the emotion inductions for anxiety and anger was reflected in an increase of self-reported state anxiety and state anger, respectively. *Second*, the anxiousness and angriness attributes of the IATs were validated by their correlations with established questionnaire measures. *Third*, in contrast to direct self-ratings, the anxiousness and the angriness IAT were not correlated with social desirability. *Fourth*, social

desirability did not moderate the correlation between direct and indirect measures. *Fifth*, the bipolar anxiousness and angriness self-ratings, as well as the observer judgments for anxiety and anger did not correlate with each other. In contrast, the anxiousness and the angriness IAT were correlated when the anxiousness IAT was the first test, $r = .49$, but not when the angriness IAT was the first test, $r = .17$. This correlation difference was marginally significant, and was attributed to a task-recoding in terms of a positive-negative self-dimension that was transferred from the anxiousness IAT onto the angriness IAT. *Sixth*, the validity of the anxiousness and the angriness IAT was marginally affected, if the test was the second rather than the first indirect test. *Seventh*, the anxiousness, but not the angriness IAT, added incremental validity over direct measures to the prediction of behavior.

5.5.2 Gender Differences

The sample in Study 2 was counterbalanced for gender. However, sex was not introduced as an independent variable in the results section because female and male participants did not differ significantly from each other with respect to the correlational analyses (Hypotheses 2-7). The only significant influence of gender was found in the anxiety induction effect (bipolar state anxiety items after the announcement of the speech minus baseline, $F_{(1, 98)} = 17.71$, $p < .001$) that was qualified by an interaction effect with gender, $F_{(1, 98)} = 5.58$, $p < .05$. Post hoc comparisons with Bonferoni correction ($p < .025$) indicated that the increase in state anxiety was true only for women, $t_{(49)} = 4.35$, $p < .001$, but not for men, $t_{(49)} = 1.41$, *n.s.*. This finding is different from Spitznagel et al. (2000) who found that women generally report more speech anxiety than man, but do not differ with respect to the increase in state anxiety. However, the studies from Spitznagel et al. used a different scale that asked for self-reports of habitual speech anxiety before, during, and after an imagined speech without a real anxiety induction. In the present study, there was no main effect of gender on any direct, indirect, or behavioral measure.

5.5.3 Behavior Prediction Through Direct and Indirect Measures

In Study 2, the observer judgments of anxious and angry behavior were predicted by the direct self-ratings. Additionally, the observer judgment of anxious behavior was predicted by the anxiousness IAT, and correlated with the duration of body adaptors and the frequency of nervous mouth movements. However, none of the behavioral anxiety and anger codings correlated significantly with either the direct measures or the IATs. The

same pattern of results was found for the shyness measures in Study 1. Thus, it is a difficult task to identify valid behavioral indicators that correlate with the observer judgments *and* direct or indirect personality self-concept measures. There might be several reasons why Study 2 failed to succeed in the search for valid behavioral cues.

Concerning the *behavioral anger* measures, the interaction between the experimenter and the participant after the computer crash was probably too short for aggregating sufficient anger indicators. The mean duration was 117 ($SD = 21$) seconds and ranged from 72 to 168 seconds. In contrast, the duration of the speech was three minutes for all participants. Importantly enough, the duration of the anger sequence did not correlate with either the observer anger judgment, direct or indirect angriness measures, or any of the anger codings (that were coded in frequencies per minute). Thus, behavioral anger measures were not confounded with the duration of the anger sequence. Nevertheless, the anger sequence was relatively short, and most behavioral anger indicators were so infrequent that even intercoder reliability was unsatisfactory. This does not, however, imply that the anger induction was inapt for the observation of angry behavior since the direct angriness measures showed predictive validity for the observer anger judgment. As many earlier attempts to study anger in the lab (e.g., Pauls & Stemmler, 2003; Wiedig, 2003), the present study only partially solved the problem that the anger sequence has to be both (a) long enough and (b) unrecognized by the participants.

Concerning the *behavioral anxiety* measures, the results from Egloff and Schmukle's (2002) Study 4 were only partially replicated in the present study. In both studies the anxiousness IAT added incremental validity over direct measures to the prediction of anxious behavior. In contrast, the observer anxiety judgment was significantly predicted by direct anxiousness measures in the present study but not in Egloff and Schmukle's Study 4. This might be due to the fact, that more direct anxiousness measures were included in the present study. The situation-specific direct measures, that is, the emotionality subscale of the Speaking Anxiety Questionnaire and the bipolar state anxiety items, were particularly strong predictors for the observer anxiety judgment in the present study. Yet, the trait form of the State Trait Anxiety Inventory also correlated marginally with the observer anxiety judgment, $r = .19$, $p < .10$, whereas this was not true for Egloff and Schmukle's study, $r = .12$, *n.s.*. However, this correlation difference was only small, and the lack of predictive validity of the direct anxiousness measure in Egloff

and Schmukle's Study 4 might also be attributed to the relatively small sample size ($N = 33$). Thus, the present study is in line with the expectation that direct measures show small to moderate validity for the prediction of behavior (Funder, 1999).

Differently from Egloff and Schmukle's Study 4, the frequency of *nervous mouth movements* was not predicted by the anxiousness IAT in the present study. This could not have been attributed to a lack of cross-lab reliability of the behavioral coding since consistency between both labs was completely satisfactory. However, the anxiousness IATs of both studies differed with respect to the attribute categories and the attribute exemplars. Attribute categories were anxiety versus calmness in Egloff and Schmukle's studies, and anxious versus self-confident in the present study. A possible post hoc explanation is that behavioral nervousness is more directly linked to a lack of calmness than to a lack of self-confidence. Therefore, the anxiety versus calmness IAT from Egloff and Schmukle might have shown better predictive validity.

Additionally, the participants in Egloff and Schmukle's Study 4 showed more nervous mouth movements than the participants in the present study, $t_{(131)} = 2.74$, $p < .01$. Possibly, behavioral anxiety was higher in Egloff and Schmukle's Study 4 due to the more evaluative nature of the speech task that asked the participants to summarize a scientific text instead of talking about euthanasia. Importantly enough, participants of both studies differed only marginally on the trait form of the State Trait Anxiety Inventory. Thus, differences between the studies should not be attributed to a general sample effect. In summary, it seems that interindividual differences in the personality self-concept of anxiousness are observed best, if one maximizes the evaluative character of the anxiety induction.

Altogether, the search for valid behavioral codings was not successful in Study 2. However, I refrained from further behavioral analysis due to the position effects and the lack of convergent and discriminant validity that were found in the angriness IAT. Nevertheless, the high correlations between the voice ratings and the observer judgments indicate that valid cues for interindividual differences in anxiousness and angriness may be found within the vocal expression of participants. Future studies of more objective vocal cues should explore this possibility.

5.5.4 Angriness, Agreeableness, Anger Expression, and Approach Behavior

Study 2 explored the implicit and explicit representations of the personality self-concept of anxiousness and angriness. Explicit representations were assessed with bipolar anxiousness and angriness self-ratings. Implicit representations were assessed by using the same words as stimuli within the IATs. The convergent validity of the bipolar anxiousness self-ratings with widespread anxiousness scales was high, $r > .70$. In contrast, the correlation between the bipolar angriness self-ratings and the trait form of the State Trait Anger Expression Inventory was only moderate, $r = .45$. This might be due to the conceptualization of anxiousness and angriness in the present study as orthogonal factors within the Big Five model of personality.

Conceptually and empirically, anxiousness versus self-confidence was strongly related to neuroticism, and unrelated to agreeableness. Angriness versus self-control was weakly related to neuroticism, and strongly related to agreeableness. In contrast, the trait form of the State Trait Anger Expression Inventory is intermediately related with emotional instability or neuroticism (Spielberger, 1988), and was also significantly correlated with all direct anxiousness measures in the present study. Differently from the trait form of the State Trait Anger Expression Inventory, the present conceptualization of angriness refers more to agreeableness and less to emotional instability or neuroticism. This may account for the moderate correlation between the bipolar angriness self-ratings and the trait form of the State Trait Anger Expression Inventory. Nevertheless, the scale was labeled as angriness because it is less broad than the Big Five dimension of agreeableness.

Alternatively, angriness versus self-control may be considered as a combination of high anger-out and low anger-control, which are strongly negatively correlated. Moreover, anger-out and anger-control show the same intermediate correlations with the trait form of the State Trait Anger Expression Inventory as the bipolar angriness self-ratings (Schwenkmezger et al., 1992). Thus, the bipolar angriness self-ratings may more directly refer to styles of anger expression than the trait form of the State Trait Anger Expression Inventory. A more direct relation to angry behavior within the bipolar angriness self-ratings is also suggested by the somewhat higher correlations with the observer anger judgment and the angry voice rating than it was obtained for the trait form of the State Trait Anger Expression Inventory (see Table 23).

Anger is a negative emotion that is related to approach behavior (see Chapter 2.6.1). In contrast, anxiety is related to avoidance behavior that is true for most of the negative emotions (e.g., sadness, disgust). Due to the relation of state anger to approach motivation, anger is associated with different EEG activation than anxiety (Harmon-Jones & Sigelman, 2001). Possibly, the automatic categorization of stimuli within the angriness IAT was somehow obstructed because angry versus self-control combines approach-related words (e.g., angry) with negative valence, and avoidance-related words (e.g., self-control) with positive valence. In contrast, avoidance-related words (e.g., anxious) are combined with negative valence, and approach-related words (e.g., self-confident) with positive valence in the anxiousness IAT. Generally, positive valence is more strongly associated with approach motivation whereas negative valence is more strongly associated with avoidance motivation (e.g., Neumann, Förster, & Strack, 2003). However, within the angriness IAT motivational direction and valence of the stimuli are inversely related. This might distort the automatic categorization of angry versus self-controlled, and further accounts for (a) the lower internal consistency within the angriness IAT (.66) than within the anxiousness IAT (.72), (b) the lack of convergent validity of the angriness IAT, and (c) the susceptibility of the angriness IAT to the transfer effect from the anxiousness IAT.

6 Study 3: Transfer Effects in Indirect Assessment

6.1 Introduction

The main purpose of the following study was to explore whether the unexpected positive correlation between the anxiousness and the angriness IAT in Study 2 was due to the salience of a positive-negative self-dimension. If participants classified the attributes within both IATs according to evaluative rather than semantic features, the transfer effect from the anxiousness IAT onto the angriness IAT can be explained. Therefore, I aimed to replicate the transfer effect of the preceding study with a different sample and to check whether the transfer effect could be blocked or strengthened using interventions that block or strengthen a positive-negative self-dimension, respectively. Additionally, I examined whether the anxiousness and the angriness IAT were correlated with a contingency-based color IAT, that assesses method-specific variance due to task-switching costs (Mierke & Klauer, in press). These research questions are discussed in the following sections.

6.1.1 Research Question 1: Interventions for Blocking and Strengthening the Transfer Effect

Study 2 provided evidence that a positive-negative self-dimension was more salient in the anxiousness IAT than in the angriness IAT. Therefore, a task-recoding in terms of a positive-negative self-dimension seemed to be more likely to occur in the anxiousness than in the angriness IAT. The task-recoding in the anxiousness IAT was assumed to cause the transfer effect from the anxiousness onto the angriness IAT. Two different interventions were examined in Study 3 in order to block the transfer effect.

One intervention employed anagrams of evaluatively neutral nouns. Participants had to identify the misplaced letters in given nouns thereby come up with the correct noun. The other intervention employed the procedure of the contingency-based IAT from Mierke and Klauer (in press). The contingency-based IAT asks for the categorization of geometrical objects and imposes an artificial contingency between the genuinely unassociated target category (color of stimuli) and the attribute category (size of stimuli) (see Chapter 2.4.2). I used meaningless strings instead of geometrical objects as stimulus material because they were easier to implement in the ERTS routines. This modification of the geometrical objects IAT was called color IAT. The detailed procedure is described in

the Methods section. The color IAT was expected to block the transfer effect because its stimuli cannot be categorized in terms of a positive-negative self-dimension.

In contrast, a positive-negative self-dimension is salient in direct self-esteem and mood measures. The transfer effect was expected to be strengthened if self-esteem and mood scales were presented between the anxiousness and the angriness IAT. Additionally, the blocking of the transfer effect through the color IAT or the anagrams seemed likely to be reversed if the self-esteem and the mood scales were presented before the angriness IAT.

6.1.2 Research Question 2: Method-Specific Variance in the IATs

Recently, Mierke and Klauer (in press) showed that method-specific variance due to task-switching can be assessed with a contingency-based IAT (see Chapter 2.4.2). As noted before, the contingency-based IAT was slightly modified for Study 3 and was used as an intervention to block the positive-negative self-dimension. The modified version, the color IAT, employed the same rationale as the geometrical objects IAT from Mierke and Klauer. It was expected that the results of the geometrical objects IAT were replicated in the color IAT. The color IAT should correlate with the absolute scores of the anxiousness and the angriness IAT when the IAT scores are calculated as conventional measures. In contrast, the color IAT should be uncorrelated with the anxiousness and the angriness IAT when the IAT scores are calculated as D measures (Greenwald et al, 2003). This would indicate that the improved D measures control for the method-specific variance that is produced by task-switching costs. Additionally, this would show that the positive correlation between the anxiousness and the angriness is not mediated by the method-specific variance that is assessed by the color IAT.

6.2 Hypotheses

Study 3 tested the following hypotheses:

Hypothesis 1 (Higher negative correlations of self-esteem and mood with direct anxiousness than with direct angriness measures).

Self-esteem and positive mood show higher negative correlations with direct anxiousness than with direct angriness measures. Therefore, a positive-negative dimension is more salient within the anxiousness IAT than within the angriness IAT. This accounts for the asymmetry of the transfer effect from the anxiousness IAT onto the angriness IAT but not vice versa.

Hypothesis 2 (Negative correlations of self-esteem and mood with the angriness IAT). When mood and self-esteem are assessed directly before the angriness IAT they show negative correlations indicating the influence of a positive-negative self-dimension on the angriness IAT.

Hypothesis 3 (Zero correlations of the D measures with the Color IAT). The improved IAT D measures do not correlate with the color IAT indicating that the correlation between the anxiousness and the angriness IAT is not mediated by the method specific variance that is assessed by the color IAT.

Hypothesis 4 (Replication of the transfer effect). The transfer effect from the anxiousness IAT onto the angriness IAT, that is, positive correlations between both IATs and a trend for the angriness IAT to correlate with direct anxiousness measures, is replicated in Study 3.

Hypothesis 5 (Intervention effects). The positive-negative self-dimension and, thus, the transfer effect is blocked through interventions that require the categorization of evaluatively neutral stimuli during an IAT or the processing of evaluatively neutral nouns.

Hypothesis 6 (Positive-negative self-dimension produces transfer effect). The transfer effect is strengthened or its blocking is reversed through self-ratings on a mood and a self-esteem scale that comprise a positive-negative self-dimension.

6.3 Methods

6.3.1 Participants and Design

180 participants were randomly assigned to the conditions of a 3 (intervention type: color IAT, anagrams, without intervention) x 2 (mood and self-esteem scale: with, without) between subjects design. Assignment was balanced for gender. Most participants were directly approached on the campus of Humboldt University, Berlin. The rest of the participants were recruited by postings at the university buildings. Participants were nonpsychology university students, native German speakers, and had not participated in the lab's previous studies. Their mean age was $M = 23.13$ years and ranged from 19 to 33 years. Participants were offered € 6 (approximately US $ 6 at the time) for taking part in a 45 minute lab experiment on personality traits.

Table 24

Overall Procedure and Design of Study 3

Cover story: Personality traits							Duration (Min.)
(a) Direct trait measures	- Trait form of the STAI and STAXI - Speaking Anxiety Scale - Bipolar self-ratings of anxiousness, angriness, conscientiousness, and intellect - Social desirability scales - Biographical data						10
(b) IAT	Anxiousness IAT						10
(c) Intervention	Color IAT		Anagrams		Without		0/5
(d) Mood and self-esteem scale	+	-	+	-	+	-	0/2
(e) IAT	Angriness IAT						10
n	30	30	30	30	30	30	~41

Note. STAI = State Trait Anxiety Inventory, STAXI = State Trait Anger Expression Inventory, IAT = Implicit Association Test, + = with, - = without.

6.3.2 Assessments and Measures

Trait measures. Trait measures were identical to Study 2 except that some scales were dropped, and the items were answered on the computer in the lab. (For more detailed information about scale formats and item numbers see Methods section of Study 2.) The questionnaire started with the trait forms of the *State Trait Anxiety Inventory STAI* (Laux et al., 1981; English version: Spielberger et al. 1970) and the *State Trait Anger Expression Inventory STAXI* (Schwenkmezger et al., 1991; English version: Spielberger, 1988). Items of both questionnaires were randomly mixed and were followed by the second series of the *Speaking Anxiety Scale* (Spitznagel et al., 2000). Next, participants had to rate their conscientiousness and intellect on 10, and their anxiousness and angriness on 5 *bipolar adjective pairs* each. Pairs were mixed in a fixed random order and presented with a trait instruction. The questionnaire concluded with the *Social Desirability Scales* by Lück and Timaeus (1969) (English version: Crowne & Marlowe, 1960) and Stöber (1999; without the Item "Have you ever consumed drugs"). Internal consistencies of all trait measures were satisfactory, $\alpha > .75$ for all scales. At the end of the questionnaire participants had to

report their age, sex, dominant hand, academic subject, length of time spent at university, whether they were still students (all were), and whether they had a permanent partner.

Mood Scale. This scale was version A of the Positive-Negative Mood Scale borrowed from the Multidimensional Comfort Questionnaire [Multidimensionaler Befindlichkeitsfragebogen, Steyer, Schwenkmezger, Notz, & Eid, 1997]. On 5-point scales (1 = not at all, 5 = very much) it assesses positive and negative mood with 2 unipolar items each (e.g., "fine"). Items were presented with a state instruction ("At the moment I feel …") and answers were coded so that higher values indicated more positive mood. Internal consistency of the Mood Scale was satisfactory, $\alpha = .88$

State Self-Esteem Scale. This scale was a short form of the State Self-Esteem Scale from Heatherton and Polivy (1991) that was translated into German by Riketta and Dauenheimer (2002). The scale deals with self-evaluations (e.g., "I feel satisfied with the way my body looks right now") that should be answered with regard to how a participant feels at the moment. Answers are given on a 5-point scale (1 = not true at all, 5 = perfectly true), with higher values indicating higher self-esteem. Out of the 20 item original scale I selected 8 items that showed corrected item-total correlations of $r > .48$ in two student samples ($N = 142$ and $N = 115$) of Riketta and Dauenheimer (personal communication, October 17, 2002). Internal consistency of the resulting scale was satisfactory, $\alpha = .80$.

Anxiousness and angriness IAT. The procedures were identical to Study 2.

Color IAT. The procedure of the color IAT was identical to the anxiousness and the angriness IAT, but the stimuli closely followed the geometrical objects IAT presented in Mierke and Klauer (in press). While target (color of stimuli) and attribute (size of stimuli) categories were equal to Mierke and Klauer, I used meaningless strings rather than geometrical objects as stimulus material. Task sequence, stimuli, and task description are depicted in Table 25. The geometrical objects IAT was developed to asses interindividual differences in task-switching performance that were shown to reliably contaminate conventional IAT measures (Mierke & Klauer, in press) but not the improved IAT D measures (Greenwald et al., 2003). The geometrical objects IAT imposes an artificial contingency between the genuinely unassociated target category (color) and attribute category (size), so that all blue stimuli are big and all red stimuli are small. I employed the color IAT in order to use a evaluatively neutral IAT procedure for studying its ability to block transfer effects between different IATs. Therefore, my procedure strictly followed the anxiousness and angriness IAT. That was also true for the aspect that (contrary to

Mierke & Klauer, in press) within the combined tasks the stimuli alternated between target and attribute. The IAT score was computed as the difference between mean response latencies in the incompatible and the compatible pairing (sequence 3 – sequence 5, see Table 25).

Table 25

Color Implicit Association Test: Task Sequence and Task Description

			Response key assignment	
Sequence	N of trials	Task	Left key	Right key
1	40	Target discrimination	Red	Blue
2	40	Attribute discrimination	Big	Small
3	80	Initial combined task	Red, big	Blue, small
4	40	Reversed target discrimination	Blue	Red
5	80	Reversed combined task	Blue, big	Red, small

Tasks	
Target discrimination: Color of strings	Attribute discrimination: Size of strings
Blue versus red	Big (22, 24) versus small (11, 12) fonts
Nonrelevant size of targets: Big (22, 24) or small (11, 12) fonts	Nonrelevant colors of attributes: Yellow, green, or pink

Note. The Color IAT imposed an artificial contingency between target (color) and attribute (size) discrimination so that all blue strings were big and all red strings were small. Strings were xyxyx, yxyxy, yxxxy, xyyyx, and xxyxx.

Anagrams. Out of a list of 800 nouns that were analyzed by M. Schwibbe, Raeder, G. Schwibbe, Borchardt, and Geiken-Pophanken (1981) I selected 35 nouns the valences of which were rated as neutral, $.08 > M > -.08$, $SD < .60$, referring to a 5-point scale ranging from -3 = negative to +3 = positive. The places of two letters were switched within each of these nouns and the nouns were presented on the screen. Participants were instructed to type in the correct noun as quick as possible. If participants did not complete the full 35 nouns within five minutes the presentation of the remaining anagrams was stopped in order to keep time comparable for all participants.

6.4 Results

6.4.1 Correlations of Direct Measures

The correlations between the state self-esteem scale and the direct anxiousness and angriness measures are depicted in the last column of Table 26. As it was expected from *Hypothesis 1*, self-esteem showed higher negative correlations with direct anxiousness than with direct angriness measures. Although correlations tended to be negative for both, they ranged from intermediate to large for direct anxiousness, and were not even marginally significant for the direct angriness measures. The correlation differences (Steiger, 1980) were nonsignificant when comparing the correlations of the bipolar anxiousness and angriness self-rating, $t_{(87)} = 1.17$, *n.s.* (one-tailed), and significant when comparing the correlations of the trait forms of the State Trait Anxiety and Anger Expression Inventories, $t_{(87)} = 2.44$, $p < .01$ (one-tailed). Although the self-esteem scale was presented with a state instruction, and anxiousness and angriness were assessed as traits, the correlational pattern with trait self-esteem might be very similar, as state and trait self-esteem are highly correlated (Heatherton & Polivy, 1991).

Concerning the correlations of the positive mood scale with direct anxiousness and angriness measures, the same pattern was true (see column 8 of Table 26). Whereas the anxiousness measures showed marginal or significant negative correlations with positive mood, the correlations of the angriness measures with positive mood were not even marginally significant. However, the correlation differences were nonsignificant when comparing the correlations of the bipolar anxiousness and angriness self-rating, $t_{(87)} = .97$, *n.s.* (one-tailed), and marginally significant when comparing the correlations of the trait forms of the State Trait Anxiety and Anger Expression Inventories, $t_{(87)} = 1.50$, $p < .10$ (one-tailed). Together, these findings illustrate that a positive-negative self-dimension was represented to a greater extent in direct anxiousness rather than in direct angriness measures. This confirmed the explanation for the transfer effect, and further demonstrated the asymmetry of the transfer effect from the anxiousness IAT onto the angriness IAT, but not vice versa.

Table 26

Reliabilities and Correlations of the Trait and State Measures in Study 3

	1	2	3	4	5	6	7	8	9
1. Bipolar anxiousness	*.83*	.06	.30***	.40***	.69***	.12	-.11	-.18^{+}	-.31***
2. Bipolar angriness		*.81*	.13^{+}	.15*	.19*	.59***	-.23**	-.04	-.15
3. Speaking Anxiety Emotionality			*.87*	.67***	.28***	.20**	-.12	.04	-.19^{+}
4. Speaking Anxiety Worries				*.86*	.40***	.28***	-.13	-.04	-.36***
5. State Trait Anxiety Inventory[a]					*.90*	.30***	-.23**	-.34***	-.59***
6. State Trait Anger Expression Inventory[a]						*.75*	-.34***	-.16	-.15
7. Social desirability							*.80*	.27**	.37***
8. Mood scale (state instruction)								*.87*	.48***
9. State self-esteem									*.80*

Note. N = 180, n = 90 for mood and self-esteem scales. Internal consistencies (Cronbach's α) are printed in *italics* along the diagonal.

[a] Trait form. $^{+}p < .05$ $*p < .05$ $**p < .01$ $***p < .001$.

Considering the other correlations in Table 26, the findings of Study 2 were clearly replicated. Again, the bipolar anxiousness and angriness self-ratings did not correlate with each other. The bipolar anxiousness self-rating correlated highly with the trait form of the State Trait Anxiety Inventory, and weakly with the trait form of the State Trait Anger Expression Inventory, whereas the opposite was true for the bipolar angriness self-rating. Finally, direct anxiousness and angriness measures tended to correlate with social desirability that was especially the case for direct angriness measures.

6.4.2 Descriptive Statistics for the Anxiousness and the Angriness IAT

Before I explore the correlations of the IATs, I will discuss briefly their descriptive statistics. The mean raw score (in milliseconds) of the *anxiousness IAT* was M = -173.4, SD = 176.5, and ranged from -748.7 to 310.0. Only 24 (13 female, 11 male) out of 180 participants had positive IAT scores. Thus, most of the participants were quicker to combine Me+self-confident and Others+anxious than the reverse mapping. The mean raw score of the *angriness IAT* was M = -153.2, SD = 124.7, and ranged from –513.6 to 123.0. Only 15 (3 female, 12 male) out of 180 participants had positive scores. Thus, most of the participants were quicker to combine Me+self-controlled and Others+angry than the

reverse mapping. The mean raw score of the *color IAT* was $M = 352.1$, $SD = 162.3$, and ranged from 66.2 to 789.4. Thus, all participants were quicker in the compatible rather than in the incompatible pairing.

Mean error rates were for the anxiousness IAT $M = 5.0\%$, $SD = 3.6\%$, for the angriness IAT $M = 3.7\%$, $SD = 3.1\%$, and for the color IAT $M = 4.1\%$, $SD = 3.5\%$. One participant had an error rate of 21.9% in the angriness IAT. Because exclusion of this participant would not affect the correlational pattern, his data were not discarded from analysis. Error rates for all other participants were below 20% in any IAT. No participant responded quicker than 300 ms in more than 10% of the trial responses in any IAT. The distributions of the D measures were not even marginally different from a normal distribution in all IATs, $Z < 1$. For every test, internal consistency was computed across the two test halves, and was acceptable for the anxiousness IAT , $\alpha = .77$, but only marginal for the angriness IAT, $\alpha = .60$, and the color IAT, $\alpha = .59$.

6.4.3 Correlations of the Anxiousness and the Angriness IAT with Self-Esteem and Mood

As it was expected from *Hypothesis 2*, the angriness IAT correlated negatively with the self-esteem and the mood scale that half of the participants ($n = 90$) completed just before the angriness IAT. Correlations were small but significant for both the self-esteem and the mood scale, $r = -.23$, $p < .05$, $r = -.21$, $p < .05$. This illustrated once more that categorization of stimuli within the angriness IAT was influenced by a positive-negative self-dimension. The anxiousness IAT, that was completed beforehand, correlated not even marginally with the self-esteem scale and marginally with the mood scale, $r = -.10$, *n.s.*, $r = -.18$, $p < .10$. As one may recall, the opposite was true for direct anxiousness and angriness measures. Only direct anxiousness measures correlated significantly with the mood and self-esteem scale whereas direct angriness measures were not even marginally correlated (see Table 26). Thus, the negative correlation of the angriness IAT was rather an indicator of its susceptibility to a positive-negative self-dimension rather than an indicator for its validity.

6.4.4 Correlations of the Anxiousness and the Angriness IAT with the Color IAT

As it was expected from *Hypothesis 3*, the D measures of the anxiousness and the angriness IAT were not correlated with the method-specific variance assessed by the color IAT, $r = -.10$, *n.s.*, $r = .07$, *n.s.*, $n = 60$. For these correlations, the absolute magnitude of the D measures was employed, and scores for the color IAT were computed on the basis of untransformed response latencies to maximize the amount of method-specific variance in such scores (Mierke & Klauer, in press). The observed zero correlations indicated that the correlation between the anxiousness and the angriness IAT was not due to their shared reliable contamination by the method-specific variance.

With regard to conventional measures, the absolute magnitude of the untransformed and the log-transformed scores of the angriness IAT correlated significantly with the color IAT, $r = .38$, $p < .01$, $r = .27$, $p < .05$. This pattern exactly replicated the findings that Mierke and Klauer (in press) obtained for an extraversion IAT and a flower-insect attitudes IAT. However, it was not true for the anxiousness IAT. For this test, the absolute magnitude of neither the untransformed scores nor the log-transformed scores correlated significantly with the color IAT, $r = .09$, *n.s.*, $r = .06$, *n.s.*. Thus, even so the conventional measures of the angriness IAT showed considerable method-specific variance due to task-switching costs, the anxiousness IAT did not. Consequently, these findings illustrated that the correlation between the anxiousness and the angriness IAT is unlikely to be mediated by the method-specific variance that was assessed by the color IAT.

The different correlations of the anxiousness IAT and the angriness IAT with the color IAT might be attributed to a position effect, since the color IAT was always presented after the anxiousness and before the angriness IAT. Nevertheless, in Mierke and Klauer's Experiment 3 (in press) an extraversion IAT correlated with method-specific variance that was assessed by a geometrical objects IAT, although the geometrical objects IAT was completed after the extraversion IAT. Therefore, it should be the subject of future studies to explore whether the method-specific variance due to task-switching increases with the number of IATs that are completed.

6.4.5 Correlations of the Anxiousness and the Angriness IAT by Intervention

Contrary to *Hypothesis 4*, the transfer effect from the anxiousness IAT onto the angriness IAT was *not* replicated when there were no intervention and no self-ratings on the mood and self-esteem scale before the angriness IAT. As it can be seen for condition (1) in Table 27 the correlation between both IATs tended to be positive but was not even marginally significant.

Table 27

Correlations of the Anxiousness and the Angriness IAT by Interventions (Study 3)

Intervention	No intervention				Color IAT				Anagrams			
Mood + SE	without		with		without		with		without		with	
Condition	(1)		(2)		(3)		(4)		(5)		(6)	
	IAT											
	Anx.	Angr.	Anx.	Angr.	Anx.	Angr.	Anx.	Angr.	Anx.	Angr.	Anx.	Angr.
Angr. IAT	.24	-	.38*	-	.46**	-	.53**	-	.38*	-	.53**	-
Bip. anx.	.54**	-.04	.24	-.03	.54**	.19	.34^{+}	.23	.46*	-.07	.35^{+}	.30
Bip. angr.	.10	<u>.42</u>*	-.23	<u>-.06</u>	.20	.04	-.08	-.18	-.25	.27	.04	.06
Emotionality	.15	-.06	.08	-.05	.19	.19	.00	.00	.27	.01	.16	.27
Worries	.35^{+}	-.13	.25	.09	.34^{+}	.15	-.07	-.01	.33^{+}	.14	.05	.21
STAI	.47**	.07	.21	.10	.50**	.24	.33^{+}	.02	.28	<u>-.22</u>	.43*	<u>.78</u>***
STAXI	.15	<u>.52</u>**	-.08	<u>-.02</u>	.14	.10	.11	-.03	.07	.47**	.29	.25
SD	-.17	-.26	-.03	.13	.32^{+}	.14	-.11	.04	.35^{+}	-.01	-.04	-.14

Note. n =30. Mood + SE = Mood and self-esteem scale (state instruction), Angr. = Angriness, Anx. = Anxiousness, Bip. anx. = Bipolar anxiousness self-rating, Bip. angr. = Bipolar angriness self-rating, Emotionality = Speaking Anxiety Emotionality, Worries = Speaking Anxiety Worries, STAI = Trait form of the State Trait Anxiety Inventory, STAXI = Trait form of the State Trait Anger Expression Inventory, SD = Social desirability.
Correlations that differed significantly due to the presentation of the mood and self-esteem scale are <u>underlined</u> ($p < .05$, one-tailed). $^{+}p < .10$ * $p < .05$ ** $p < .01$ *** $p < .001$.

More importantly, the angriness IAT showed sizeable convergent validity with the bipolar angriness self-rating and the trait form of the State Trait Anger Expression

Inventory. Inspection of the scatterplots revealed homogenous distributions. Thus, the angriness IAT did not correlate significantly with either the anxiousness IAT or with direct anxiousness measures. Therefore, Hypothesis 4 was not confirmed.

With regard to *Hypothesis 5*, I explored the groups that completed the color IAT or the anagrams but not the mood and self-esteem scales before the angriness IAT. Contrary to Hypothesis 5, the *color IAT* that required a categorization of senseless strings according to color and size was not sufficient to block the transfer effect. As it can be seen for condition (3) in Table 27, the anxiousness and angriness IAT correlated considerably in this group. The angriness IAT did not correlate with direct angriness but showed small positive correlations with direct anxiousness measures that nonetheless failed to reach significance. In contrast, in the *anagram* group, the angriness IAT tended to show weaker correlations with the anxiousness IAT, and instead correlated with direct angriness measures. As it is shown for condition (5) in Table 27, the angriness IAT correlated weakly, but due to the small sample size nonsignificantly, with the bipolar angriness self-rating, and intermediately with the trait form of the State Trait Anger Expression Inventory. However, the angriness IAT still correlated significantly with the anxiousness IAT. Thus, Hypothesis 5 was confirmed, demonstrating that the anagrams increased the convergent validity of the angriness IAT. Nevertheless, the anagrams were incapable of entirely eliminating the correlation between the anxiousness and the angriness IAT. On the other side, the color IAT was generally inappropriate to block the transfer effect

Concerning the effects of the mood and self-esteem scale that should promote a positive-negative self-dimension, *Hypothesis 6* was successfully confirmed. When these scales were presented between the anxiousness and the angriness IAT, the convergent validity of the angriness IAT was diminished, and both IATs tended to correlate more highly with each other. Although this correlation increase was not significant, the pattern was replicated in all three groups (see condition (2), (4), and (6) in Table 27). More importantly, the decrease of convergent validity with direct angriness measures was significant for the no intervention group (see condition (2) in Table 27). In addition, there was a significant increase in the correlation between the angriness IAT and the trait form of the State Trait Anxiety Inventory in the anagram group (see condition (6) in Table 27). It should be noted that the latter correlation showed a homogenous scatterplot and was not driven by outliers. Altogether, the transfer effect was clearly strengthened through the presentation of the mood and self-esteem scales.

6.5 Discussion

This section first summarizes the main findings of Study 3. Then I discuss why the transfer effect might not have been replicated in the no-intervention group and refer to the problem of small sample sizes.

6.5.1 Summary of the Main Findings

Study 3 explored whether the unexpected positive correlation between the anxiousness and the angriness IAT that was found in Study 2 was caused by a task-recoding in terms of a positive-negative self-dimension. Therefore, the salience of the positive-negative self-dimension was manipulated, and the effects on the correlations of the angriness IAT were studied. Study 3 tested six hypotheses.

First, direct anxiousness measures showed a pattern of stronger correlations with negative self-esteem and negative mood than direct angriness measures. Thus, a positive-negative self-dimension was more salient in the stimuli of the anxiousness IAT than of the angriness IAT. This explains the asymmetry of the transfer effect from the anxiousness IAT onto the angriness IAT. *Second*, the angriness IAT correlated significantly with both, negative self-esteem and negative mood, if they were presented directly before the angriness IAT. This indicated that a positive-negative dimension influenced the angriness IAT. *Third*, in contrast to the conventional scores, the improved D measure of the angriness IAT did not correlate with the method-specific variance that was assessed by the color IAT. *Fourth*, unexpectedly, in the no-intervention group, the transfer effect from the anxiousness IAT onto the angriness IAT was not replicated and the IATs were only weakly correlated. This lack of replication is discussed in more detail in Chapter 6.5.2. *Fifth*, only the anagrams but not the color IAT were capable of reducing the transfer effect and securing the convergent validity of the angriness IAT. However, even the anagrams did not entirely eliminate the correlation between the anxiousness and the angriness IAT. *Sixth*, when a positive-negative self-dimension was made salient through the presentation of self-esteem and mood scales, the transfer effect was strengthened. This was apparent from a pattern of higher correlations between the anxiousness and the angriness IAT, and from a lack of convergent and discriminant validity of the angriness IAT.

6.5.2 Lack of Replication of the Transfer Effect in the No-Intervention Group

Surprisingly, the transfer effect from the anxiousness IAT onto the angriness IAT failed to be replicated in the no-intervention group of Study 3. The transfer effect was explained by the salience of a positive-negative self-dimension. However, there may be three reasons why a positive-negative self-dimension was weaker in Study 3 than in Study 2.

First, in Study 2 direct anxiousness and angriness measures were completed at home within one week before the lab experiment. In contrast, in Study 3, direct anxiousness and angriness measures were completed during the lab experiment and before the anxiousness and the angriness IAT. Previous studies showed that correlations between IATs and direct measures are affected by the order in which direct measures and the IATs are presented (Bosson et al., 2000), and that correlations between IATs and direct measures are not affected by the presentation order (Nosek et al., 2003). However, a recent meta-analysis (Hofmann, Gawronski, et al., 2003) suggests that first administering the direct measures increases correlations between IATs and direct measures. More importantly, it is theoretically plausible that direct-indirect correlations increase when the direct measures are completed first, because this makes the existing associations more accessible (Fazio, 1995; Strack & Deutsch, in press). It is possible that the angriness IAT in Study 3 was more robust against the transfer effect from the anxiousness IAT because the presentation of the direct angriness measures made the implicit self-concept of angriness more accessible. The presentation of the direct measures before the IATs seemed to have had at least some effect, as direct-indirect correlations tended to be generally higher in Study 3 than in Study 2. For instance, the overall correlation of the anxiousness IAT with the bipolar anxiousness self-ratings tended to be higher in Study 3 ($r = .39$, $p < .001$, $N = 180$) than in Study 2 ($r = .25$, $p < .05$, $N = 100$). Similarly, the correlation of the angriness IAT with the bipolar angriness self-ratings tended to be higher in Study 3 ($r = .23$, $p < .05$, $n = 90$) than in Study 2 ($r = .11$, *n.s.*, $N = 100$).

Second, the cover story in Study 3 was “personality traits”, whereas in Study 2 it was “concentration and personality”, since the cover story in Study 2 had to be discreet in regards to the emotion inductions. Especially during the lab experiment of Study 2, the focus was on concentration tests. Therefore, some participants may have categorized anxious versus self-confident (within the anxiousness IAT) and angry versus self-

controlled (within the angriness IAT) in terms of distracted versus concentrated. This might have strengthened the transfer effect. It could also explain why the direct-indirect correlations in Study 2 were, on the whole, lower than in Study 3. Nevertheless, a task-recoding in terms of distracted versus concentrated is unlikely to explain the transfer effect alone because the transfer effect was asymmetrical.

Third, in Study 2 participants completed an optimistic risk questionnaire directly before the IATs. In this questionnaire participants rated the probability of positive (e.g., "I married someone wealthy.") and negative (e.g., "I had a heart attack before age 50.") events during their lifetime. Items were recoded such that high scores indicated optimism. The optimistic risk questionnaire did not correlate with either the anxiousness IAT, the angriness IAT, or explicit angriness measures. Yet, it showed significant negative correlations with all explicit anxiousness measures ($r < -.25$). Therefore, the optimistic risk questionnaire might have made a positive-negative dimension in Study 2 more salient, and, thus, strengthened the transfer effect from the anxiousness onto the angriness IAT.

Altogether, it is clear that correlations between the IATs and direct measures are dependent on the context (cf. the special issue of the Journal of Personality and Social Psychology, 71, 2001). Particularly, direct-indirect correlations seem to be affected by the aspects of the personality self-concept that are activated before the IATs. In Study 3, a positive-negative self-dimension was probably weakened by presenting the direct angriness measures beforehand, by focusing on personality traits, and by omitting the optimistic risk questionnaire. Therefore, the transfer effect might have not been replicated in the no intervention group. Nevertheless, it may be assumed that the evaluatively neutral anagrams still contributed to the reduction of the transfer effect in the anagram group for two reasons. First, the transfer effect was replicated in the group that completed the color IAT. Second, the transfer effect was strengthened by the presentation of the self-esteem and the mood scales, that made, in contrast to the evaluatively neutral anagrams, a positive-negative self-dimension more salient.

6.5.3 Small Sample Sizes

The aim of Study 3 was to test correlational hypotheses about blocking and strengthening of the transfer effect from the anxiousness IAT onto the angriness IAT. The sample sizes for the 6 experimental conditions were quite small (n = .30). It was expected by Hypothesis 5 that the neutral anagrams, as well as the color IAT would block the transfer effect. Results showed that this was true only for the former but not the latter. Therefore, these groups could not be pooled. Additionally, the transfer effect was not replicated in the no-intervention group, and the correlations of the anxiousness and the angriness IAT varied considerably among the groups that completed the self-esteem and the mood scale before the angriness IAT (see Table 27). Consequently, the experimental conditions were discussed individually.

In any correlational study, large sample sizes, at least $N \geq 50$, but better $N \geq 100$, are called for. Otherwise, results may be driven by a few uncharacteristic participants who are unidentifiable as outliers. I looked for outliers who might have distorted the correlations, and I examined all of direct and indirect measures in every experimental condition individually. However, this search was unsuccessful. Similarly, the inspection of the scatterplots revealed that the correlations showed homogenous distributions. Still, the importance of large sample sizes might be illustrated by the fact that the correlation between the anxiousness IAT and the bipolar anxiousness self-ratings varied considerably between the conditions, from r = .24 to r = .54. This was true although the experimental conditions were identical up to the presentation of the anxiousness IAT. Thus, results of small samples such as in the present study should be considered with caution.

Nevertheless, a pattern of three important results became evident in all experimental conditions. First, in contrast to the bipolar anxiousness and angriness self-ratings, the correlation between the anxiousness IAT and the angriness IAT was positive and different from zero. Second, the correlation between the IATs was higher if a positive-negative self-dimension was made salient through the self-esteem and the mood scales. Third, the convergent and the discriminant validity of the angriness IAT with direct self-report measures was distorted by the presentation of the self-esteem and the mood scales.

More importantly, a different study with a larger sample size (N = 97) replicated the position effect on the angriness IAT that was found in Study 2 (Teige, Schnabel, Banse, & Asendorpf, in press). In that study, the sequence of the shyness and the angriness IAT was

counterbalanced across participants. The correlations between the shyness and the angriness IAT tended to be higher when the shyness IAT was completed as first rather than second test, $r = .34$ versus $r = -.01$, $z = -1.73$, $p < .10$ (two-tailed). Similar to the anxiousness IAT, the shyness IAT may have made a positive-negative self-dimension more salient, because shyness and anxiousness were highly correlated concerning direct self-reports ($r = .69$). The study provided further evidence that the transfer effect on the angriness IAT is (a) asymmetrical and (b) most likely caused by the salience of a positive-negative self-dimension.

7 General Discussion

The present research comprised 3 studies. In *Study 1*, the shyness IAT and a new indirect procedure, the shyness IAP, showed considerable convergent validity. Both of the indirect measures were much less susceptible to faking instructions than the direct self-ratings. Additionally, under faking instructions, the correlations of direct and indirect measures with shy behavior decreased more strongly for the direct rather than for the indirect measures. However, there was a lack of valid behavioral codings for controlled and spontaneous shy behavior. Therefore, the double dissociation pattern of Asendorpf et al.'s (2002) Study 1 was not replicated. In *Study 2*, the anxiousness IAT added incremental validity over direct anxiousness measures to the prediction of anxious behavior. However, the angriness IAT was affected by a transfer effect from the anxiousness IAT. *Study 3* provided further evidence that this transfer effect was due to the salience of a positive-negative self-dimension.

The General Discussion refers to three aspects of these findings. First, to what extent are indirect measures influenced by the semantic meaning or by the positive and negative valence of the stimuli? Second, why are direct and indirect measures of the personality self-concept different from each other? Third, what can be learned from the present and other findings for the assessment of the implicit personality self-concept?

7.1 Semantic Meaning versus Valence

The findings of Study 2 and Study 3 suggested that the positive correlation between the anxiousness IAT and the angriness IAT was attributable to a task-recoding in terms of a positive-negative self-dimension. This raises a question about the extent to which IAT measures are driven by the semantic meaning as opposed to the positive or negative valence of the stimuli. If IATs mainly reflect the ease with which one combines positive versus negative stimuli with Me, then the IATs represent self-esteem IATs (e.g., Greenwald & Farnham, 2000) rather than indirect measures of different personality traits. Can the empirical findings of different self-concept IATs be re-interpreted in terms of implicit measures of self-esteem?

Concerning the anxiousness IAT, the answer might be 'yes'. In the studies by Egloff and Schmukle (2002), the anxiousness IAT predicted performance decrement due to failures in a concentration test, and anxious behavior during an evaluative speech task. Both behaviors may also be predicted by a 'pure' self-esteem IAT that does not directly

refer to anxiousness (cf. Greenwald & Farnham, 2000). The results from the anxiousness IAT in Study 2 of the present research may be re-interpreted using the same logic. The same reasoning can be applied to the shyness IAT (Asendorpf et al., 2002; Study 1 of the present research) as well, such that shy behavior could be predicted by low self-esteem. Already at the level of direct measures, shyness and anxiousness are negatively correlated with self-esteem (Cheek & Melchior, 1990; Judge, Erez, Bono, & Thoresen, 2002). Thus, it is difficult to disentangle valence and semantic meaning in anxiousness and shyness because a valid portion of these traits already contains negative valence.

There are at least three indicators that the semantic meaning of the stimuli may influence IAT scores. *First*, self-esteem IATs show that most people are quicker in combining Me with positive attributes than in combining Me with negative attributes. Thus, most individuals have positive implicit self-esteem (Greenwald & Farnham, 2000). This pattern did not hold to be true in the shyness IAT of Asendorpf et al.'s (2002) Study 1. The shyness IAT indicated that about 40 % of the participants are more shy than non-shy. If the shyness IAT could be re-interpreted as self-esteem IAT, one would expect fewer participants to show positive scores in the shyness IAT.

Second, the research on priming procedures provides evidence that semantic and affective priming procedures show effects even with very short SOA (stimulus onset asynchrony, i.e., the interval between start of prime and start of target stimulus) (Greenwald, Draine, & Abrams, 1996; Klauer & Musch, 2003). Thus, both semantic and evaluative information is processed very quickly. Importantly, affective priming seems to depend on whether respondents have to identify evaluative or non-evaluative target attributes. Affective priming does usually not occur when targets are classified on the basis of non-evaluative features. (De Houwer et al., 2002, Klauer & Musch, 2003). To my knowledge, the present evidence of affective and semantic priming does not indicate that priming effects are more influenced by either valence or semantic meaning of stimuli. In contrast, the characteristics of the priming task seem to influence whether valence information or semantic information causes priming effects.

Third, in two experimental groups of Study 3, the angriness IAT showed convergent validity for direct angriness measures (up to $r = .52$) in addition to discriminant validity for direct anxiousness measures (all $r < .14$). This finding could not have been explained had the participants classified stimuli within the angriness IAT only in terms of a positive-negative dimension.

On the other hand, task-recoding in terms of a positive-negative self-dimension cannot be completely ruled out in self-concept IATs due to a relatively strong connection between the concept of self and positive valence (Greenwald et al., 2002). Thus, the categorization of Me versus Others may automatically activate the positive-negative dimension. Self-concept IATs with *two* positive or *two* negative traits as target attributes could block the positive-negative dimension. For instance, one could employ a self-concept IAT with anxiousness versus angriness as attribute categories. However, this IAT should, then, show identical results for those individuals who score high on anxiousness and angriness and those who score low on both traits. The results of recent studies provide evidence that it is difficult to separate the IATs' measure of relative associations into two independent measures (Nosek et al., 2003).

The positive-negative self-dimension may alternatively be weakened by the following factors. *First*, the positive or negative valence of IAT stimuli and IAT attribute categories should not be too extreme (cf. Footnote 4 from De Houwer, 2001; Steffens et al., 2003). For instance, although shyness is a negatively valenced trait, it is only so to a moderate extent. Likewise, with regard to anxiousness and angriness, one should consider stimuli with more positive valence, for example, cautious and resistant, respectively. *Second*, one should make an effort to highlight the semantic meaning of stimuli and to block a positive-negative self-dimension. The presentation of evaluatively neutral anagrams seemed to be successful in Study 3.

7.2 Dissociations between Direct and Indirect Measures of the Personality Self-Concept

There would be no interest in researching indirect measures if indirect and direct assessment procedures measured identical constructs. In order to differentiate between operationalizations and constructs, in this work, the terms *direct* and *indirect* measures were used to label procedures, and the terms explicit and implicit representations were used to label the constructs. Similarly, there may be two sources for dissociations between direct and indirect measures of the personality self-concept: (a) theoretically-based dissociations between explicit and implicit representations at the construct level; (b) method factors in direct and indirect measures at the assessment level.

Concerning the construct level, explicit representations were regarded as propositional categorizations within the *Reflective System*, and implicit representations as

associative clusters within the *Impulsive System* of the Reflective-Impulsive Model from Strack and Deutsch (in press) (see Chapter 2.2). Thus, explicit representations should be better predictors of controlled behavior, and implicit representations should be better predictors of spontaneous behavior. Recently, Asendorpf et al. (2002) carried out a double dissociation procedure between the explicit and implicit personality self-concept of shyness. A direct shyness questionnaire uniquely predicted controlled (but not spontaneous) shy behavior, whereas a shyness IAT uniquely predicted spontaneous (but not controlled) shy behavior. However, the results of the present studies, Study 1 and Study 2, showed that it is difficult to differentiate between indicators of spontaneous and controlled behavior.

Attempts to show predictive validity of indirect measures often follows an incremental validation strategy. This means that the studies usually explore whether indirect measures predict variance in relevant criteria in addition to direct measures of the same construct (for a review, cf. Fazio & Olson, 2003). In Study 2 of the present research, the anxiousness IAT added incremental validity over direct self-ratings to the prediction of the observer anxiety judgments. Conceptually, the incremental validity of the indirect measures might be attributed to two differences between explicit and implicit representations, (a) implicit representations have more direct access to the associative store than explicit representations, (b) explicit representations might be biased due to social desirability concerns (cf. Chapter 2.3).

Biases based on social desirability also affect *method factors* of direct and indirect measures. For instance, whereas direct self-ratings are certainly *fakable* (cf. Study 1), there is a controversy about IATs being fakable or not (cf. Chapter 2.4.1). Study 1 provided evidence that the shyness IAT was to some extent fakable by. However, effects were much smaller than for direct self-ratings.

Another methodical issue could refer to the question of whether the indirect procedure employs *negation* or not. For instance, the typical target categories of self-concept IATs are Me versus Others. For the IAP in Study 1, target categories were Me versus Notme. According to Strack and Deutsch (in press), the Impulsive System is not able to negate information. More precisely, the Impulsive System is not able to assign a true or a false value to the relation between two concepts. Instead, the Impulsive System only connects or does not connect concepts using episodic and semantic links that are available within the associative store. Therefore, it may be an interesting topic for further

research to explore whether indirect measures that employ negation (e.g., the GNATs "Go/No-Go Association Tasks", Nosek & Banaji, 2001; the shyness IAP of Study 1) are influenced by the Reflective System more than indirect measures that do not employ negation. However, the shyness IAP employed in Study 1 did not seem to be more controllable than the shyness IAT since the IAP was even less susceptible to faking instructions. Additionally, the IAT and the IAP contained a negation for the attribute category, that is, Shy versus Nonshy.

Another characteristic of indirect measures is that they can be influenced by the category frame of the categorization task *and* by individual stimulus features (cf. Fazio & Olson, 2003). In contrast, direct self-ratings, for example, bipolar adjectives, are judged individually, that is, they are only influenced by individual stimulus features. To obtain the mean scale, the bipolar items are combined such that every item is weighed equally. Concerning IATs, there is some evidence that individual stimulus features have an effect on the IAT score, while the category frame is more influential (cf. Chapter 2.4.2).

Finally, dissociations between direct and indirect measures may also be caused by the *context dependency* in indirect measures. Although indirect measures seem to be not affected by emotion inductions (Schmukle & Egloff, 2003; cf. Study 1 of the present research), they were shown to be influenced by other contextual variables (cf. Mitchell et al., 2003). Importantly enough, the present research revealed evidence that self-concept IATs are affected by the salience of a positive-negative self-dimension (see Chapter 7.2).

In summary, dissociations between direct and indirect self-concept measures might be attributed to differences at the construct and at the measurement level. At the construct level, implicit representations differ from explicit representations because of their more direct access to the associative store and a more direct effect on spontaneous behavior. At the measurement level, there are method factors that are characteristic of indirect procedures rather than of direct procedures. Indirect procedures are less fakable, presumably less apt to assess negated concepts, are influenced by the category frame and stimulus features, and are susceptible to contextual variables, particularly to the salience of a positive-negative self-dimension.

7.3 Recommendations for Future Research

This section summarizes some aspects of indirect measures, particularly of IATs, that may help future research on the implicit personality self-concept to be more successful. *First*, IAT D measures (Greenwald et al., 2003) were shown to control for method-specific variance due to task-switching (cf. Hypothesis 3 from Study 3). Therefore, future data reduction procedures should employ D measures.

Second, it would be most useful to standardize the IAT procedure, that is, using a standardized number of trials in the five different IAT tasks, and even using a standardized number of stimulus items per IAT category. Although the effects of such variations may be insignificant (Nosek et al, 2003), it could make research on IATs much more comparable. Particularly, it was shown that the effect of task order (IAT effects are larger if the compatible pairing is completed before the incompatible pairing) was reduced by increasing trials in the reversed target discrimination task (see Table 3). A number of 40 trials (like in the present studies) seems to be optimal for the reversed target discrimination (Nosek et al., 2003). This finding is important for the assessment of the personality self-concept because the task order effect seems to maximize interindividual differences for participants with positive IAT scores (see Chapter 2.4.2)

Third, the order of the compatible and incompatible IAT pairing should not be counterbalanced for the assessment of the personality self-concept. Otherwise, order variance is confounded with interindividual variance. *Fourth*, correlations between direct measures and the IAT seem to be higher if the direct measures are completed before the IATs (Hofmann, Gawronski, et al., 2003; cf. Chapter 6.5.2; for findings revealing no order effect, cf. Nosek et al., 2003). Thus, if one aims to maximize consistency between direct and indirect measures, one should apply direct measures first.

Fifth, results of the present studies indicate that IATs may be influenced by the semantic meaning and by the positive or negative valence of IAT stimuli. Therefore, personality self-concept IATs should try to use moderately valenced stimuli and to block the salience of a positive-negative self-dimension.

Sixth, there are promising results for self-concept IATs that use stimuli that describe behaviors (e.g., fistcuff) rather than personality traits (e.g., aggressive) (Banse & Fischer, 2002). Within the Reflective-Impulsive Model from Strack and Deutsch (in press), behavioral motor schemata are subsumed into the Impulsive System. Thus, it seems

plausible that stimuli describing behavior are more strongly represented in the associative clusters of the Impulsive System rather than abstract personality traits. Therefore, IATs with behavior describing items may provide a more direct access to the associative store.

Seventh, one should include the stimuli of the indirect measures as direct self-ratings. This allows for a fair comparison between direct and indirect measures. Additionally, the internal consistency of the stimuli that represent one trait can also be checked at the direct level. This could indicate whether the stimuli fit well in the superordinate category.

Eighth, if one employs more than one IAT in order to assess different traits within one study, one should take into consideration that transfer effects between the IATs might distort the IAT effects. Study 2 and Study 3 provided evidence that the position effect on the angriness IAT was due to a task-recoding in terms of a positive-negative self-dimension that was transferred from the anxiousness IAT. A positive-negative task-recoding might be quite salient in self-concept IATs due to the strong connection of Me with positive valence (Greenwald et al., 2002). In addition, there may be other superordinate categories for other traits, for example, male-female, young-old, intelligent-unintelligent, that can bias the IAT scores.

8 Discover *Id*: The Conclusion

When Freud (1923) named the deep, inaccessible part of personality the id, this was certainly an original term. Employing a model from more scientific-empirical Social Cognition research, I differentiated between explicit and implicit representations of one's own personality and considered them as elements of reflective and impulsive information processing, respectively (Strack & Deutsch, in press). Using the traits of shyness, anxiousness, and angriness as examples, I assessed implicit representations of the personality self-concept with the Implicit Association Tests (IATs, Greenwald et al., 1998) and the new Implicit Association Procedures (IAPs) as the tools for indirect measures. In contrast to direct questionnaire measures that assess the explicit personality self-concept, indirect measures are chronometric procedures that avoid asking direct self-judgment questions.

The results showed four important dissociations between direct and indirect measures in the assessment of the personality self-concept. First, indirect measures were more robust against faking than direct measures. Second, the convergent validity between indirect measures was lower than that between direct measures. Third, indirect measures added incremental validity to the prediction of behavior. Fourth, indirect measures were less apt for the concurrent assessment of two traits within one sample than direct measures.

The latter factor was explained by the fact that indirect measures are influenced by the semantic meaning and the positive versus negative valence of stimuli. Whether the former or the latter most likely affected the results depended on whether a positive-negative self-dimension was made salient. The angriness IAT was particularly distorted by the salience of a positive-negative self-dimension. This may be explained by the fact that angriness, though negatively valenced, is related to approach behavior, whereas negative valence is usually associated with avoidance behavior (Neumann et al., 2003). Therefore, a positive-negative self-dimension might have been weaker in the angriness IAT, and the angriness IAT was more affected by a positive-negative self-dimension in the context.

In my opinion, an important aspect of present indirect measures, such as the IATs, is that they allow participants to refuse cooperation. I consider this aspect a justification for further research rather than a deficiency, because the results of such research cannot be employed against the will of examinees. From my point of view, the main purpose of indirect measures is not that they may circumvent the self-presentational strategies of respondents, but that indirect measures lead to a better understanding of the information processes that underlie implicit and explicit representations.

In conclusion, the unresolved issue of semantic meaning and valence being confounded, and the relatively low convergent validity between indirect measures provided evidence that indirect measures are not yet ready to be used as standard instruments for personality assessment. On the other hand, the development of indirect measures such as the IATs (Greenwald et al., 1998) represents a ground-breaking work for two reasons. First, IAT measures assess interindividual differences with internal consistencies that are satisfactory and much higher than the internal consistencies of other indirect procedures, for instance, priming measures (e.g., Banse, 2001; Bosson et al., 2000; Kawakami & Dovidio, 2001). Second, in the present and in other studies, IAT measures were shown to increase the prediction of behavior (e.g. Asendorpf et al., 2002; Egloff & Schmukle, 2002; McConnell & Liebold, 2001). Indirect measures, even in their infancy, are an indispensable research instrument to assess implicit representations of the personality self-concept in order to draw a more holistic picture of personality.

9 References

Aiken, L. R. (1996). *Personality assessment: Methods and practices (2nd ed.).* Seattle: Hogrefe & Huber.

Asendorpf, J. B. (1989a). Shyness as a Final Common Pathway for Two Different Kinds of Inhibition. *Journal of Personality and Social Psychology, 57,* 481-492.

Asendorpf, J. B. (1989b). *Soziale Gehemmtheit und ihre Entwicklung [Social inhibition and its development]*. Berlin: Springer.

Asendorpf, J. B. (1992). Beyond stability: Predicting inter-individual differences in intra-individual change. *European Journal of Personality, 6,* 103-117.

Asendorpf, J. B., Banse, R., & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology, 83,* 380-393.

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

Banse, R., & Fischer, I. (2002, November). *Implizite und explizite Aggressivität, soziale Erwünschtheit und die Vorhersage aggressiven Verhaltens. [Implicit and explicit aggression, social desirability, and the prediction of aggressive behavior]*. Paper presented at the 7th Workshop Aggression, Universität Wien.

Banse, R., & Gawronski, B. (2003). Die Skala Motivation zu vorurteilsfreiem Verhalten: Skaleneigenschaften und Validierung. [The scale Motivation to Act Without Prejudice: Psychometric properties and validity]. *Diagnostica, 49,* 4-13.

Banse, R., Seise, J., & Zerbes, N. (2001). Implicit attitudes towards homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift für Experimentelle Psychologie, 48,* 145-160.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology, 71,* 230-244.

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation

effect: Unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology, 32,* 104-128.

Behringer, J. (1994). *Experimental Run Time System (Version 3.18).* Frankfurt, Germany: BeriSoft Corporation.

Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology, 81,* 828-841.

Bless, H., Fiedler, K., & Strack, F. (2003). *Social cognition: How individuals construct social reality.* Hove, UK: Psychology Press.

Blümke, M., & Friese, M. (2003). *Do irrelevant features of stimuli influence IAT effects?* Manuscript submitted for publication.

Bosson, J. K., Swann, W. B., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology, 79,* 631-643.

Brauer, M., Wasel, W., & Niedenthal, P. (2000). Implicit and explicit components of prejudice. *Review of General Psychology, 4,* 79-101.

Brendl, C. M., Markman, A. B., & Messner, C. (2001). How do indirect measures of evaluation work? Evaluationg the inference of prejudice in the implicit association test. *Journal of Personality and Social Psychology, 81,* 760-773.

Brendl, C. M., Markman, A. B., & Messner, C. (2003). *Indirectly measuring evaluations of several attitude objects in relation to a neutral reference point.* Manuscript submitted for publication.

Brickenkamp, R. (1994). *Test d2 Aufmerksamkeits-Belastungs-Test (8th ed.) [Test d2 attention stress test].* Göttingen: Hogrefe.

Buchner, A., & Wippich, W. (2000). On the reliability of implicit and explicit measures. *Cognitive Psychology, 30,* 227-259.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81-105.

Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology.* New York: Guilford.

Cheek, J. M., & Melchior, L. A. (1990). Shyness, self-esteem, and self-consciousness. In H. Leitenberg (Ed.), *Handbook of social and evaluation anxiety* (pp. 47-82). New York: Plenum Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.).* Hillsdale, NJ: Erlbaum.

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

Costa, P. T., Jr. & McCrae, R. R. (1992). *Revised NEO Personality Inventory and NEO Five Factor Inventory professional manual.* Odessa, FL: Psychological Assessment Resources.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24,* 349-354.

Crozier, W. R. (2001). *Understanding shyness*. Basingstoke: Palgrave.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 12,* 163-170.

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology, 81,* 800-814.

Dasgupta, N., McGhee, D. E., Greenwald, A. G., & Banaji, M. R. (2000). Automatic preference for white Americans: Eliminating the familiarity explanation. *Journal of Experimental Social Psychology, 36,* 316-328.

De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. *Journal of Experimental Social Psychology, 37,* 443-451.

De Houwer, J. (2003a). The Extrinsic Affective Simon Task. *Experimental Psychology, 50,* 77-85.

De Houwer, J. (2003b). A structural analysis of indirect measures of attitudes. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 219-244). Mahwah, NJ: Lawrence Erlbaum.

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

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology, 82,* 62-68.

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an Implicit Association Test for assessing anxiety. *Journal of Personality and Social Psychology, 83,* 1441-1455.

Egloff, B., & Schmukle, S. C. (2003). Does social desirability moderate the relationship between implicit and explicit anxiety measures? *Personality and Individual Differences, 35,* 1697-1706.

Egloff, B., Schwerdtfeger, A., & Schmukle, S. C. (2003). *One-year stability of the Implicit Association Test-Anxiety.* Manuscript submitted for publication.

Ekman, P., & Friesen, W. V. (1972). Hand movements. *Journal of Communication, 22,* 353-374.

Ekman, P., & Friesen, W. V. (1978). *The facial action coding system.* Palo Alto, CA: Consulting Psychologists Press.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist, 49,* 709-724.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. In M. P. Zanna (Ed.), *Advances in experimental social psychology (Vol. 23,* pp. 75-109). San Diego, CA: Academic Press.

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

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology, 54,* 297-327.

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.

Fiedler, K., & Blümke, M. (2003). *Faking the IAT: Aided and unaided response control on the Implicit Association Tests.* Manuscript submitted for publication.

Fiedler, K., Messner, C., & Bluemke, M. (2003). *Unresolved problems with the "I", the "A" and the "T": Logical and psychometric critique of the Implicit Association Test*

(IAT). Unpublished manuscript.

Fiske, D. W. (1987). On understanding our methods and their effects. *Diagnostica, 33,* 188-194.

Freud, S. (1923). *Das Ich und das Es [The ego and the id]*. Leipzig: Internationaler Psychoanalytischer Verlag.

Friesen, W. V., & Ekman, P. (1984). *EMFACS-7.* San Francisco: Unpublished report, Human Interaction Laboratory, University of California, San Francisco.

Funder, D. C. (1999). *Personality judgment: A realistic approach to person perception.* San Diego, CA: Academic Press.

Garner, W. R. (1969). Speed of discrimination with redundant stimulus attributes. *Perception & Psychophysics, 6,* 221-224.

Gawronski, B. (2002). What does the Implicit Association Test measure? A test of the convergent and discriminant validity of prejudice related IATs. *Experimental Psychology, 49,* 171-180.

Gawronski, B., Ehrenberg, K., Banse, R., Zukova, J., & Klauer, K. C. (2003). It's in the mind of the beholder: The impact of stereotypic associations on category-based and individuating impression formation. *Journal of Experimental Social Psychology, 39,* 16-30.

George, C., Kaplan, N., & Main, M. (1985). *The Adult Attachment Interview.* Unpublished manuscript. University of California, Berkley.

Govan, C. L., & Williams, K. D. (in press). Changing the affective valence of the stimulus items influences the IAT by redefining the category labels. *Journal of Experimental Social Psychology.*

Gray, J. (1987). Perspectives on anxiety and impulsivity. A commentary. *Journal of Research in Personality, 21,* 493-509.

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

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellot, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109,* 3-25.

Greenwald, A. G., Draine, S. C., & Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science, 273,* 1699-1702.

Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology, 79,* 1022-1038.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85,* 197-216.

Greenwald, A. G., McGhee, D.E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74,* 1464-1480.

Harmon-Jones, E., & Sigelman, J. (2001). State anger and prefrontal brain activity: Evidence that insult-related relative left-prefrontal activation is associated with experienced anger and aggression. *Journal of Personality and Social Psychology, 80,* 797-803.

Heatherton, T. F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology, 60,* 895-910.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Hodapp, V. (1991). Das Prüfungsängstlichkeitsinventar TAI-G: Eine erweiterte und modifizierte Version mit vier Komponenten [The Test Anxiety Inventory TAI-G: An expanded and modified version with four components]. *Zeitschrift für Pädagogische Psychologie, 5,* 121-130.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2003). *A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures.* Manuscript submitted for publication.

Hofmann, W., Gschwendner, T., & Schmitt, M. (2003). *On the moderation of the implicit-explicit consistency: The role of introspection and adjustment.* Manuscript submitted for publication.

John, O. P., & Srivastava, S. (2001). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of*

personality: Theory and research (2nd ed., pp. 102-138). New York: Guilford.

Jones, W. H., Briggs, S. R., & Smith, T. G. (1986). Shyness: Conceptualization and measurement. *Journal of Personality and Social Psychology, 51,* 629-639.

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are Measures of Self-Esteem, Neuroticism, Locus of Control, and Generalized Self-Efficacy Indicators of a Common Core Construct? *Journal of Personality and Social Psychology, 83,* 693–710.

Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology, 81,* 774-788.

Kawakami, K., & Dovidio, J. F. (2001). The reliability of implicit stereotyping. *Personality and Social Psychology Bulletin, 27,* 212-225.

Kim, D-Y. (2003). After the South and North Korea Summit: Malleability of explicit and implicit national attitudes of South Koreans. *Peace and Conflict: Journal of Peace Psychology, 9,* 159-170.

Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and theories. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 219-244). Mahwah, NJ: Lawrence Erlbaum.

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

Krohne, H. W., Egloff, B., Kohlmann, C. W., & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der "Positive and Negative Affect Schedule" (PANAS) [Investigations with a German version of the Positive and Negative Affect Schedule (PANAS)]. *Diagnostica, 42,* 139-156.

Laux, L., Glanzmann, P, Schaffner, P., & Spielberger, C.D. (1981). *Das State-Trait-Angstinventar [The State-Trait Anxiety Inventory].* Weinheim: Beltz.

Lazarus, R. S. (1991). *Emotion and adaptation.* New York: Oxford University Press.

Leary, M. R. (1983). Social anxiousness: The construct and its measurement. *Journal of Personality Assessment, 47,* 66-75.

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social*

Psychology, 81, 146-159.

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology, 81,* 842-855.

Lück, H. E., & Timaeus, E. (1969). Skalen zur Messung Manifester Angst (MAS) und sozialer Wünschbarkeit (SDS-E und SDS-CM) [Scales for the assessment of manifest anxiety (MAS) and social desirability (SDS-E and SDS-CM)]. *Diagnostica, 15,* 134-141.

McConnell, A. R., & Liebold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37,* 435-442.

Mierke, J., & Klauer, K. C. (2001). Implicit association measurement with the IAT: Evidence for effects of executive control processes. *Zeitschrift für Experimentelle Psychologie, 48,* 107-122.

Mierke, J., & Klauer, K. C. (in press). Method-specific variance in the Implicit Association Test. *Journal of Personality and Social Psychology*.

Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General, 132,* 455-469.

Murray, H. A. (1943). *Thematic Apperception Test. Manual.* Cambridge, MA: Harvard University Press.

Neely, J. (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. 264-337). Hillsdale, NJ: Erlbaum.

Neumann, R., Ebert, M., Gabel, B., Guelsdorff, J., Kranich, H., Lauterbach, C., & Wiedl, K. (1998). Vorurteile zwischen Bayern und Norddeutschen: Die Anwendung einer neuen Methode zur Erfassung evaluativer Assoziationen in Vorurteilen [Prejudice prevailing between Bavarians and North Germans: Application of a new method for the assessment of evaluative associations in prejudices]. *Zeitschrift für Experimentelle Psychologie, 45,* 99-108.

Neumann, R., Förster, J., & Strack, F. (2003). Motor compatibility: The bidirectional link between behavior and evaluation. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation. Affective Processes in Cognition and Emotion* (pp. 762-768). Mahwah,

NJ: Lawrence Erlbaum.

Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition, 19*, 625-666.

Nosek, B. A., & Banaji, M. R. (2002). (At least) two factors moderate the relationship between implicit and explicit attitudes. In I. R. K. Ohme & M. Jarymowicz (Eds.), *Natura Automatyzmow* (pp. 49-56). Warszawa: WIP PAN & SWPS.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therfore math // me. *Journal of Personality and Social Psychology, 8,* 44-59.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2003). *Understanding and using the Implicit Association Test: II. Methodological issues.* Unpublished manuscript.

Ones, D. S., & Viswesvaran, C. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance, 11,* 245-269.

Ostendorf, F. (1990). *Sprache und Persönlichkeitsstruktur: Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit [Language and personality structure: Validity of the five-factor model of personality)]*. Regensburg, Germany: S. Roderer.

Ottaway, S. A., Hayden, D. C., & Oakes, M. A. (2001). Implicit attitudes and racism: Effects of word familiarity and frequency in the Implicit Association Test. *Social Cognition, 19,* 97-144.

Paulhus, D. L. (1984). Two-component model of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598-609.

Pauls, A. C., & Stemmler, G. (2003). Repressive and defensive coping during fear and anger. *Emotion, 3,* 284-302.

Pervin, L. A., & John, O. P. (Eds.). (2001). *Handbook of personality: Theory and research* (*2nd ed*.). New York: Guilford.

Plessner, H., Wänke, M., Friese, M., & Haar, T. (2003). *Measuring implicit consumer attitudes and predicting brand choice.* Manuscript submitted for publication.

Riketta, M., & Dauenheimer, D. (2002). *Unpublished German translation of the State Self Esteem Scale from Heatherton and Polivy (1991).* Mannheim: University of Mannheim.

Rothermund, K., & Wentura, D. (2001). Figure-ground asymmetries in the Implicit Association Test (IAT). *Zeitschrift für Experimentelle Psychologie, 48,* 94-106.

Rothermund, K., & Wentura, D. (2003). *The Implicit Association Test (IAT) as a new "window to the mind" - but what do we see behind the pane?* Manuscript submitted for publication.

Rudman, L. A., Ashmore, R. D., & Gary, M. L. (2001). "Unlearning" automatic biases: The malleability of implicit prejudice and stereotypes. *Journal of Personality and Social Psychology, 81,* 856-868.

Rudman, L. A., Greenwald, A. G., Mellot, D. S., & Schwartz, J. L. K. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition, 17,* 437-465.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 13,* 501-518.

Schmukle, S. C., & Egloff, B. (2003). *Does the Implicit Association test for assessing anxiety measure trait and state variance?* Manuscript submitted for publication.

Schwenkmezger, P., Hodapp, V. & Spielberger, C. D. (1992). *Das State-Trait-Ärgerausdrucks-Inventar STAXI [The State-Trait-Anger-Expression-Inventory STAXI]*. Göttingen: Huber.

Schwibbe, M., Räder, K., Schwibbe, G., Borchardt, M., & Geiken-Pophanken, G. (1994). Zum emotionalen Gehalt von Substantiven, Adjektiven und Verben [The emotional content of nouns, adjectives, and verbs]. In W. Hager & M. Hasselhorn (Eds.), *Handbuch deutschsprachiger Wortnormen [Handbook of German word norms]*. Göttingen: Hogrefe.

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

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

Smith, E. R., & Mackie, D. M. (2000). *Social Psychology* (*2nd ed.*). Philadelphia, PA: Psychology Press.

Spielberger, C. D. (1980). *Test Anxiety Inventory ("Test Attitude Inventory"). Preliminary*

professional manual. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D. (1988). *State-Trait-Anger-Expression-Inventory (STAXI). Research Edition.* Odessa, FL: Psychological Assessment Resources.

Spielberger, C. D., Auerbach, S. M., Wadsworth, A. O., Dunn, T. M., & Taulbee, E. S. (1973). Emotional reactions to surgery. *Journal of Consulting and Clinical Psychology, 40,* 33-38.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *STAI. Manual for the State-Trait Anxiety Inventory.* Palo Alto, CA: Consulting Psychologists Press.

Spitznagel, A., Schlutt, S., & Schmidt-Atzert, L. (2000). Redeängstlichkeit: Entwicklung eines mehrdimensionalen Verfahrens [Public speaking anxiety: Construction of a multidimensional inventory]. *Diagnostica, 46,* 47-55.

Steffens, M. C., & Buchner, A. (2003). Implicit Association Test: Separating transsituationally stable and variable components of attitudes towards gay men. *Experimental Psychology, 50,* 33-48.

Steffens, M. C., Jelenec, P., Anheuser, J., Georgens, K., Hülsebusch, T., Lichau, J., & Still, Y. (2003). *Automatically evaluating individuals and groups: A two-factor account of reaction time differences in the Implicit Association Test.* Manuscript submitted for publication.

Steffens, M. C., & Plewe, I. (2001). Items' cross-category associations as a confounding factor in the Implicit Association Test. *Zeitschrift für Experimentelle Psychologie, 48,* 123-134.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87,* 245-251.

Stemmler, G., Heldmann, M., Pauls, C. A., & Scherer, T. (2001). Constraints for emotion specificity in fear and anger: The context counts. *Psychophysiology, 38,* 275-291.

Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. (1997). *Der Mehrdimensionale Befindlichkeitsfragebogen (MDBF) [The Multidimensional Comfort Questionnaire].* Göttingen: Hogrefe.

Stöber, J. (1999). Die Soziale-Erwünschtheits-Skala-17 (SES-17): Entwicklung und erste Befunde zu Reliabilität und Validität [The social desirability scale 17: Development and first results on reliability and validity]. *Diagnostica, 45,* 173-177.

Strack, F., & Deutsch, R. (in press). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review.*

Taylor, J. A. (1953). A personality scale of manifest anxiety. *Journal of abnormal social Psychology, 48,* 285-290.

Teige, S., Schnabel, K., Banse, R., & Asendorpf, J. B. (in press). Assessment of multiple implicit self-concept dimensions using the Extrinsic Affective Simon Task (EAST). *European Journal of Personality.*

van Ijzendoorn, M. H. (1995). Adult attachment representations, parental responsiveness, and infant attachment: A meta-analysis on the predictive validity of the Adult Attachment Interview. *Psychological Bulletin, 117,* 387-403.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063-1070.

Wells, G. L., & Petty, R. E. (1980). The effects of overt head movements on persuasion: Compatibility and incompatibility of responses. *Basic and Applied Social Psychology, 1,* 219-230.

Wiedig, M. (2003). *Ärger und soziale Regeln. Ärgererleben und Ärgerverhalten in der Selbst- und Fremdwahrnehmung [Anger and social rules. Anger experience and anger behavior in self-ratings and observer ratings].* Unpublished dissertation thesis, Enrst-Moritz-Arndt-Universität Greifswald.

Wigboldus, D. (2003). *Single target implicit association tests.* Paper presented at the 3rd Workshop on Implicit Diagnostics. Bonn: University of Bonn.

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

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology, 81,* 815-827.

Wyer, R. S., & Srull, T. K. (Eds.). (1994). *Handbook of social cognition.* Hillsdale, NJ: Erlbaum.

10 Appendix

10.1 German IAT Stimuli

Ich	Andere	Schüchtern	Nichtschüchtern
selbst	andere	gehemmt	ungezwungen
meine	eure	unsicher	sicher
eigen	fremd	zaghaft	wagemutig
mir	euch	zurückhaltend	freimütig
ich	ihr	verschlossen	kontaktfreudig

Ängstlich	Sicher	Ärgerlich	Kontrolliert
ängstlich	sicher	ärgerlich	kontrolliert
furchtsam	wagemutig	aufbrausend	bedächtig
unsicher	selbstvertrauend	unbeherrscht	selbstbeherrscht
besorgt	sorglos	hitzköpfig	fügsam
übervorsichtig	unbeschwert	motzig	friedlich

10.2 Experimental Set-Up and Screen Design of the IAP in Pilot Study 2 and in Study 1