Investigating the attitude towards ambiguity: Interindividual differences in automatic activations of evaluations of ambiguity

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Raphael Titt

aus Remscheid

Tübingen

2021

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	29.07.2021
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	apl. Prof. Dr. René Ziegler
2. Berichterstatter:	Prof. Dr. Barbara Kaup

Danksagung

Danke René, dass Du mich immer so gut beraten hast, obwohl ich manchmal den Anschein erweckt habe „beratungsresistent“ zu sein, was sich tatsächlich nur auf ganz wenige nebensächliche Aspekte wie z.B. kontraintuitive Kodierungen bezieht. Keine Sorge, es sollten eigentlich keine mehr in der Arbeit enthalten sein. Bei allen relevanten Punkten war ich immer sehr dankbar über Deine kritischen Fragen, Anmerkungen und Ratschläge, die mein wissenschaftliches Arbeiten und damit auch diese Arbeit weit vorangebracht haben.

Danke Mandy, dass auch Du mich immer so gut beraten hast. Leider habe ich Dich erst spät im Verlauf der Arbeit, aber noch rechtzeitig genug, gebeten meine zweite Betreuerin zu sein. Durch dein Wissen hat die Arbeit noch einmal eine andere Dynamik bekommen und ich hätte gerne noch mehr Studien mit Dir und René als Team geplant, gerne natürlich auch welche, die hypothesenkonformen Ergebnisse geliefert hätten, aber dann wäre es vielleicht nicht so spannend gewesen.

Was hätte ich ohne die emotionale Unterstützung von Dir Vénice nur gemacht? Du hast mir immer zugehört, hast mir immer Mut und Kraft gegeben weiter zu machen, egal wie düster ich gelegentlich zu Denken pflegte. Dafür danke ich Dir von ganzem Herzen. Egal wie enttäuscht ich über wirre Daten, genervt über abstürzende Programme oder wie verzweifelt ich über den Berg von unerledigten Dingen auch war, wenn ich euch, Eleana und Marelie, am Abend gesehen habe, kamen automatisch so starke positive Gefühle und Gedanken, die alles andere relativieren; dann weiß man wieder was einem das Wichtigste im Leben ist. Danke Eleana, dass es Dir immer wichtig war, dass ich auch die „Uni-Maus“ mitnehme. Als wüsstest Du mit deinen knapp drei Jahren, wie gut es mir tat sie zu sehen und nicht anders konnte als an Dich zu denken.

Natürlich danke ich auch allen 3632 Teilnehmerinnen und Teilnehmern für die Teilnahme an den Studien, die in diese Arbeit eingeflossen sind. Ohne eure Daten wäre diese Forschung offensichtlich nicht möglich gewesen.

Table of Contents

Abstract

Direct measures of (in)tolerance of ambiguity provided evidence for a variation in liking of ambiguity. Given the limitations of these measures, we developed a direct measure of attitude towards ambiguity. However, the main part of this thesis deals with the questions, whether there are interindividual differences in the automatic activations of evaluations of subjective ambiguity (single information triggers multiple distinct representations). We developed a database with norms for ambiguous and unambiguous German words. A subset of these words matched for several dimensions was used in two indirect measures. The Implicit Association Test (IAT) assessed the relative strength of associations of ambiguity and clearness with positive and negative valence, respectively. In the Evaluative Priming (EP) paradigm ambiguous and unambiguous primes preceded targets with positive or negative valence. This allows to draw inference about the automatic evaluations of the primes. In order to validate potential variances in the automatic evaluations of the primes (indicated by interindividual differences in latencies as a function of prime type and target type), information about the attitude towards ambiguity was used as a moderator. Apart from the first study, which investigated the relation of direct measures with the IAT and showed unrelatedness, all other studies used the EP paradigm, using either the IAT score (studies 2 – 4) or the induced associations of opposite valence with ambiguity and clearness (studies 5 – 6) as moderators. Studies 2 and 3 provided evidence for interindividual differences in the activation of evaluations of ambiguity via a three-way interaction of prime type, target type and IAT score. However, this three-way interaction was not found in the replication study 4. In study 5, there was an interaction of prime type, target type, and induction but opposite to the expected direction. The post-hoc explanation for the partial contrast effect was further investigated by manipulating the SOA in study 6. However, in this study the induction had no influence on latencies in the EP paradigm, but there was a prime type, target type and SOA interaction. This can be explained in terms of contrast (long SOA) and assimilation (short SOA) effects if we consider the evidence across the aforementioned studies showing that, on average, participants had a stronger association of ambiguity with negative valence and clearness with positive valence. Summarized, evidence for interindividual different automatic evaluations of ambiguity was weak, but the results of the EP paradigms indicated a more negative (or less positive) automatic evaluation of ambiguity compared to clearness. The implications of automatic evaluations of the mental representation of ambiguity are discussed.

Zusammenfassung

Direkte Maße der Ambiguitäts(in)toleranz lieferten Hinweise darauf, dass die Bewertungen gegenüber Ambiguität variieren. Angesichts der Einschränkungen dieser Maße entwickelten wir ein direktes Maß für die Einstellung gegenüber Ambiguität. Der Hauptteil dieser Arbeit beschäftigt sich jedoch mit der Frage, ob es interindividuelle Unterschiede in den automatischen Aktivierungen von Bewertungen subjektiver Ambiguität gibt (eine Information löst mehrere unterschiedliche Repräsentationen aus). Wir entwickelten eine Datenbank mit Normen für mehrdeutige und eindeutige deutsche Wörter. Eine Teilmenge dieser Wörter, die hinsichtlich mehrerer Dimensionen gematcht waren, wurde in zwei indirekten Maßen verwendet. Der implizite Assoziationstest (IAT) bewertete die relative Stärke der Assoziationen von Ambiguität und Klarheit mit positiver bzw. negativer Valenz. Im Evaluative Priming (EP) Paradigma gingen ambige und eindeutige Primes Targets mit positiver oder negativer Valenz voraus. Mit diesem Paradigma können Rückschlüsse auf die automatische Evaluation der Primes gezogen werden. Um die potenzielle Varianz bei der automatischen Evaluation der Primes (angezeigt durch interindividuell unterschiedliche Latenzen als Funktion der Primes und der Targets) zu validieren wurden Informationen über die individuelle Einstellung zur Ambiguität als Moderator verwendet. Außer der ersten Studie, in der die Beziehung direkter Maße zum IAT untersucht wurde und keine Beziehung gefunden wurde, verwendeten alle Studien das EP Paradigma. Dabei diente entweder der IAT-Score (Studie 2 – 4) oder die Induktion von Assoziationen entgegengesetzter Valenz von Ambiguität und Klarheit (Studie 5 – 6) als Moderator. Die Studien 2 und 3 lieferten Evidenz für die interindividuell unterschiedliche Aktivierung von Evaluationen von Ambiguität durch die dreifache Interaktion von Prime, Target und IAT-Score. Diese Dreifach-Interaktion wurde jedoch in der Replikationsstudie 4 nicht gefunden. In Studie 5 gab es eine Interaktion von Induktion, Prime und Target, jedoch entgegen der erwarteten Richtung. Die post-hoc-Erklärung für den partiellen Kontrasteffekt wurde durch Manipulation der SOA in Studie 6 weiter untersucht. In dieser Studie hatte die Induktion jedoch keinen Einfluss auf die Latenzen im EP-Paradigma. Es gab jedoch eine Interaktion von Prime, Target und SOA. Dies kann durch Kontrasteffekte (bei langer SOA) und Assimilationseffekte (bei kurzer SOA) erklärt werden, wenn die Evidenz aus den oben genannten Studien berücksichtigt wird. Sie zeigten, dass die Teilnehmenden im Mittel eine stärkere Assoziation von Ambiguität mit negativer Valenz und Klarheit mit positiver Valenz hatten. Insgesamt fanden wir nur schwache Evidenz für interindividuelle Unterschiede bei der automatischen Aktivierung von Evaluationen von Ambiguität. Im Allgemeinen sprechen die Ergebnisse des EP-Paradigmas jedoch für eine negativere (oder weniger positive) automatische Evaluation von Ambiguität im Vergleich zu Klarheit. Die Implikationen der automatischen Evaluation der mentalen Repräsentation von Ambiguität werden diskutiert.

1 Introduction

Allport (1935) wrote about attitudes that they are “our methods for finding our way about in an ambiguous universe” (p. 806). When he used the term *ambiguous,* he did not refer to ambiguity as a specific phenomenon but rather referred to the amount and complexity of information to which we are exposed. In this sense “attitudes simplify our day-to-day existence, enabling efficient appraisal of the objects that we encounter” (Fazio, 2007, p. 629). However, there are true ambiguities all over our ambiguous universe, too. But whether or not there is a specific attitude determining our evaluative responses towards ambiguous stimuli in the space of this universe of attitudes, is unclear. This is even more surprising when looking at the existing literature: the concept of *intolerance of ambiguity* was introduced more than 70 years ago by Frenkel-Brunswik (1949). However, there is a mismatch between our understanding of ambiguity and the usual conceptualization of the phenomenon in literature (see chapter 2.3.1). As “science without definitions of basic constructs would be chaotic” (Eagly & Chaiken, 2007, p. 583), it is necessary to clarify precisely what we mean by ambiguity. In our view, ambiguity occurs if a piece of information triggers multiple but a finite number of *distinct* interpretations or more generally speaking representations (for the discrimination between vagueness and ambiguity see Black, 1937). In contrast, the term *ambiguous* is often used to refer to situations of otherwise unclear information (see chapter 2.1). Another divergence from the literature is that we consider the concept intolerance towards ambiguity as an attitude instead of a personality trait (see chapter 2.3). This attitudinal conceptualization, compared to the personality view, highlights the role of *evaluative* responses to ambiguity and presupposes a narrower definition of situations causing the response. Furthermore, treating ambiguity as an attitude object implies the existence of evaluations (or an evaluative process based on evaluatively useful information) that can vary inter- and intraindividually. In order to make inferences about the existence of an attitude, it is necessary to investigate the evaluations of an object. Investigating evaluations can only be done by assessing observable evaluative responses which can be done in manifold ways (for a short overview of possible different evaluative responses see De Houwer, 2009). We are particularly interested in the more *automatic* evaluative responses to ambiguity. In doing so, we do not assume that this approach reveals more “true” evaluations than evaluative responses derived from approaches which allow for deliberate information processing. However, as deliberate processing probably leads to disambiguation (selecting one meaning), we argue that the investigation of evaluations of the core of the phenomenon of ambiguity (the mental state of having multiple distinct meanings in mind) can be better done with paradigms which reduce the possibility deliberate processing. Most importantly, although there are plenty of direct measures of (in)tolerance towards ambiguity (see chapter 2.3.1), direct empirical evidence for interindividually different evaluations due to the processing of ambiguous information corresponding to the directly assessed attitude towards ambiguity are missing. Based on the existing literature, the question remains if there is indeed a negative (positive) evaluation triggered by an ambiguous *stimulus* for a person who agrees (disagrees) with the statement "I don't like ambiguity". In that sense it is unclear if the directly assessed (evaluative) response to an item dealing with the phenomenon of ambiguity is indicative for the evaluative response to an ambiguous stimulus. We are trying to answer this question while developing a direct measure of attitude towards ambiguity (see chapter 3.2).

However, the main part of this work focuses on the automatic evaluations of ambiguity, especially, whether there is evidence for this type of evaluative response at all and whether there are interindividual differences of these automatic evaluations, in the sense that there is substantial interindividual variation in liking of ambiguity. When we use the term automatic, we refer to the characteristics of unintentionality and fastness (for information on the concept of automaticitiy see Bargh, 1994; Moors, 2016). In the general discussion, we will shed more light on the automaticity features relevant for our paradigms. We used two well-known indirect measures of attitudes to investigate the interindividual differences in automatic activations of evaluations of ambiguity (to be more precise: to the mental state of having two or more distinct representations in mind derived from one stimulus): the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) and the Evaluative Priming [EP] Paradigm (Fazio, Jackson, Dunton, & Williams, 1995; Fazio, Sanbonmatsu, Powell, & Kardes, 1986), always together within one sample (see Studies 2 to 4). Additionally, we also tried to manipulate the attitude towards the concept of ambiguity and observed whether the EP paradigm was sensitive to this manipulation (see Studies 5 and 6). This indirect approach (the investigation of a potential moderation of the EP effect by the IAT score or the manipulation of attitude towards ambiguity) allows to infer whether an automatic evaluation of ambiguity occurred and more importantly (which is implied by the literature on intolerance of ambiguity) whether these evaluations triggered by an ambiguous stimulus differ interindividually (more than random variation). The reason why we think that it is necessary to use two indirect measures or a manipulation prior to the EP paradigm in order to answer these questions and especially why we selected the EP paradigm and the IAT will be explained in chapter 2.3.2.

As nearly all kinds of stimuli (e.g., sounds, phrases, pictures, words) have the potential to activate multiple distinct representations which are linked to the meaning of that stimulus, it is important to select the type of stimulus which allows for the best investigation of the relevant research question. As our central question is about the existence of (interindividually different) automatic evaluations of the mental state of having multiple vs. a single meaning in mind, we selected lexically (un)ambiguous words as stimuli for several reasons: (1) There is evidence that the meanings are activated fast, unintentionally and (more or less) simultaneously (see chapter 2.2). As the automaticity of the meaning activation is a necessary condition (in our paradigms) in order to investigate the automatic evaluation of the state of having multiple vs. one meaning in mind, we will give an overview of experiments designed to investigate this automaticity of activation of multiple meanings from lexically ambiguous words in chapter 2.2. Especially the short presentation times (which can be used for single words) are important in some paradigms to reduce or eliminate voluntary control and strategic influence (Fiedler & Hütter, 2014). (2) Automatic evaluations are highly sensitive to the context and the individual stimulus properties (see Blair, 2002). For example, Wittenbrink, Judd, and Park (2001) showed that the indirectly assessed attitude towards black people changes as a function of the valence of the context (family barbecue vs. gang incident) in which the attitude object is presented. Instead of pictures, sentences or passages which carry a lot of additional information, single words are parsimonious in terms of context and potentially reduce the activation of unintended representations which have the power to override the evaluation of the core of the phenomenon. Within the framework of the development of direct scales of ambiguity it will be investigated whether different evaluative responses arise depending on the context in which ambiguity occurs (see chapter 3.2). Presenting single words that are neutral in terms of valence, which are selected to differ solely in the dimension of ambiguity seems to be the best way of dealing with the high context sensitivity of automatic evaluations. (3) Lexically ambiguous and unambiguous words can be easily matched for several properties like valence, arousal, abstractness, frequency in written language, and number of letters, thereby excluding several potential confounding variables. As no German word list with ratings for these dimensions especially designed for matching ambiguous and unambiguous words was available, we created our own database (chapter 3.1), which served as a basis for stimulus selection for all experiments reported in this thesis.

2 Theoretical Background

2.1 Ambiguity

Ambiguity is an attractive topic in psychological research: a PsychInfo search with the keyword “Ambiguity” offers more than 4000 results. Unfortunately, in psychological research the term ambiguity is used for labeling different entities. Reviewing the usage of the term ambiguity between 1933 and 1970, Norton (1975) ascertained eight different uses of ambiguity. Only 28 percent of the authors used the term to refer to multiple meanings of a stimulus. Others used it to refer to vague, incomplete, fragmented, unstructured, inconsistent or contradictory stimuli, to probabilistic situations or used it synonymously with unclearness and uncertainty. The uses of meanings differ between disciplines. For example, in psychologic-economic research ambiguity is usually treated as subjective (un)certainty of estimation of probabilities and this (un)certainty depends, for instance, on the reliability and quantity of information (Ellsberg, 1961). Vagueness, an ambiguity related but distinct phenomenon can also cause uncertainty. Although there is philosophical (e.g., Black, 1937; Gullvåg & Næss, 1996) and linguistic (e.g., Kennedy, 2011) research comparing ambiguity and vagueness, in psychological research the terms are often used interchangeably (see, Norton, 1975). A key feature of vague expressions (e.g., small, dimly) is the lack of sharp boundaries (Keefe, 2000). A vague stimulus therefore offers indefinite different interpretations, an ambiguous stimulus offers only finite multiple distinct interpretations (Black, 1937). The paper by Norton (1975) points out that psychologists dealing with ambiguity investigate different phenomena. Therefore, as mentioned when we use the term ambiguity, we refer to the phenomenon that a stimulus triggers more than one distinct interpretation. According to our view on ambiguity, a stimulus has to elicit multiple interpretations within one person. It is not sufficient that a stimulus triggers multiple interpretations across individuals, which is often regarded as a criterion for ambiguity (Norton, 1975).

In other disciplines the term ambiguity refers only to a very special phenomenon. The physicist Caglioti (1992) defined ambiguity as “coexistence, at a critical point, of two aspects or schemes of reality which are mutually exclusive and which have become physically observable” (p. 17). Only one of Norton’s (1975) eight categories characterizing the multiple uses of ambiguity is compatible with this definition, namely, ambiguity as multiple meanings of a stimulus. In contrast to Caglioti (1992), he distinguished between ambiguity as an objective property of the stimulus (objective ambiguity) and ambiguity arising due to having multiple interpretations of one stimulus in mind (subjective ambiguity). From a psychological point of view the distinction between objective and subjective ambiguity is important, especially when the goal is the investigation of the cognitive, affective or behavioral outcomes resulting from processing ambiguity. We can assume that in most cases effects of an objectively ambiguous stimulus differ depending on the state of subjective ambiguity. If an objectively ambiguous stimulus does not trigger different interpretations and hence triggers no subjective ambiguity, it probably does not lead to different effects compared to an objectively unambiguous stimulus. The importance of subjectivity of ambiguity in psychological research is unique; considering, for example, the prominent view in literature that “ambiguity is a fact in the text-a double system of mutually exclusive clues.” (p. 12) (Rimmon, 1977) or in linguistics where much research is done regarding the objective (micro)structure of a text constituting ambiguity (Bauer, Knape, Koch, & Winkler, 2010).

Exemplarily, let us turn to some (social) psychological research dealing with the phenomenon of ambiguity (for an overview see, Ziegler, 2010). In social psychology the effect of (biased) interpretation of ambiguous information depending on accessibility of other information is a matter of interest. For instance, in the classic experiment by Higgins, Rholes, and Jones (1977) participants were unobtrusively primed with evaluative polarized trait concepts prior to reading a passage about a person (Donald) whose ambiguous (past and planned) behaviors (e.g. “cross the Atlantic in a sailboat”) could be interpreted in a positive (“adventurous”) or negative (“reckless”) way. The evaluations and characterizations of Donald given afterwards were in line with the evaluative tone of the primed traits, however only for semantically applicable trait concepts. It should be noted that under special conditions semantically nonapplicable traits as primes can influence the interpretation of ambiguous information, too (see, Croizet & Fiske, 2000; Stapel & Koomen, 2000, 2005). On the one hand, one could regard the operationalization of ambiguity (ambiguous behavior) used in this paradigm as an instance of objective ambiguity, because it allows for two distinct interpretations; on the other hand, this operationalization in combination with the manipulation of accessibility of information challenges the existence of subjective ambiguity (having both interpretations in mind).

This biased interpretation of ambiguous information can be found for subliminally primed trait concepts, too (Bargh & Pietromonaco, 1982; Erdley & D'Agostino, 1988). But not only primed trait concepts seem to have the potential to influence the interpretation of ambiguous behavior. Rudman and Lee (2002) showed that violent and misogynistic rap music as a prime influences the interpretation of the ambiguous behavior of black men. Moreover, the effect of biased interpretation of ambiguity depending on primed information is not limited to long term priming but extends to short term priming, too (for a comparison between short and long term priming see, Wentura & Rothermund, 2014). For example, there is evidence that automatic evaluations of subliminal, positive or negative non-trait prime words influence the interpretation of ambiguity, established via semantically unrelated homographs (homonyms with same spelling) with evaluative positive and negative meanings (Ferguson, Bargh, & Nayak, 2005, Exp. 1). This study additionally showed that not only ambiguity established via ambiguous behavior (like in the Donald paradigm) is sensitive to biased interpretation. Furthermore, the effect of biased interpretation is not limited to primed concepts but generalized to dispositions, too. There is evidence for a biased (threat-related) interpretation or disambiguation of homophones (Eysenck, MacLeod, & Mathews, 1987) and ambiguous sentences (Eysenck, Mogg, May, Richards, & Mathews, 1991; MacLeod & Cohen, 1993) depending on the degree of trait anxiety. In a similar vein, the level of depression seems to determine the negativity of interpretation of ambiguity (Lawson, MacLeod, & Hammond, 2002). Likewise, the amount of worries seems to determine the degree of threat-related interpretation of ambiguous pictures although no different psychophysiological reactivity was found (Kirschner, Hilbert, Hoyer, Lueken, & Beesdo-Baum, 2016). Nonetheless, not only clinically relevant traits have the potential to bias the interpretation of ambiguity. Interindividual differences in procrastination, conscientiousness and extroversion can also determine the interpretation of ambiguous sentences about time (S. E. Duffy & Feist, 2014). Furthermore, coping styles (e.g. vigilance or cognitive avoidance) seem to have an impact on interpretation, encoding and retrieval of ambiguous information (Hock & Krohne, 2004). Hugenberg and Bodenhausen (2004) showed that the degree of racial prejudice biases the interpretation (speeded racial categorization) of racial ambiguous faces when the facial expression (hostility) is in line with the racial prejudice (for a replication see, Hutchings & Haddock, 2008). Even motivational states have an influence on the interpretation of ambiguity. Balcetis and Dunning (2006) provided evidence that motivated reasoning (wishful thinking) biases the process of visual perception preconsciously such that people more often see and report only the interpretations of ambiguous figures that correspond to their preferences. Even the presentation of ambiguous stimuli (Chinese character) in the right or left visual field bias the evaluation in such a manner that the stimuli presented in the left (compared to the right) visual field are evaluated more positively (Koch, Holland, & van Knippenberg, 2009). This was explained by the view of specialized hemispheres in terms of affective processing.

While the focus of the research mentioned so far is to provide evidence for biased interpretation of ambiguous information depending on primed concepts, individual differences in traits or motivational states et cetera, whereas in persuasion literature this sensitivity of ambiguous information for biased interpretation is used for example in order to investigate the co-occurrence and interplay of heuristic and systematic processes in persuasion (Bohner, Chaiken, & Hunyadi, 1994). A common operationalization of ambiguity in that domain is the mixture of strong and weak arguments within one message (see Chaiken & Maheswaran, 1994) or the formulation of arguments of moderate strength (Ziegler & Diehl, 2003). Both approaches for operationalizing of ambiguity do not seem to trigger multiple distinct interpretations in a direct way (objective ambiguity) although the presented information could be interpreted in different ways depending on other variables (subjective ambiguity).

In the saying-is-believing paradigm (Higgins & Rholes, 1978) ambiguity plays a central role, too. This paradigm revealed that formulating a biased message about ambiguous information about a person in line with the receiver's attitude towards that person results in biased memory of that information for the sender. Echterhoff, Higgins, and Levine (2009) argue that ambiguity leads to a feeling of uncertainty which can be reduced by adapting the message to the known attitude of the receiver. Pierucci, Echterhoff, Marchal, and Klein (2014) investigated the role of ambiguity via manipulating the amount of ambiguity (known or unknown outcome) for the same story in a between-subjects design. Biased message formulation and biased memory of the information in the story (in line with the receiver's attitude) only occured in the ambiguous condition. Ambiguity in the saying-is-believing paradigm is established either by descriptions of a single ambiguous behavior (similar to the Donald paradigm, see above) or by mixing evaluative conflicting behavioral descriptions (Pierucci et al., 2014). In the latter cases one could argue that it is more a manipulation of ambivalence as it probably results in positive and negative associations regarding a single attitude object (see, Conner & Sparks, 2002; Scott, 1966).

Although ambiguity is used in different paradigms and different fields and obviously inspired research and led to new insights, there are only few studies dealing with the affective consequences of ambiguity. In the literature dealing with stigmatized groups, the affective consequences of attributional ambiguity play an important role (see, Crocker, Cornwell, & Major, 1993; Crocker, Voelkl, Testa, & Major, 1991). Attributional ambiguity means that a person from a stigmatized group can interpret feedback from an outgroup member in terms of reflecting a true and unbiased evaluation or a biased evaluation in a negative (prejudiced) or positive (e.g. socially desirable) way. The focus of this research is not on the affective consequences of being in a state of attributional ambiguity but on the affective consequences of applying an (biased vs. unbiased) interpretation to the feedback.

This short excursion of ambiguity in psychological research demonstrates on the one hand that the interpretation/disambiguation of ambiguity depends on currently activated information (e.g. primed traits, automatic evaluations, stereotypes) and on the other hand that there are several possible ways to manipulate the (un)ambiguousness of stimuli or situations. Therefore, in order to investigate the affective reaction to the state of having multiple distinct interpretations or meanings in mind (subjective ambiguity), we have to consider potential disambiguation due to activated information and face the problem that there are several ways of establishing ambiguity, too. Furthermore, there is the question of how to assess an (automatic) evaluative reaction to the experience of subjective ambiguity (see chapter 2.3.2). One approach to establishing subjective ambiguity is to use verbal information. However, ambiguity can occur at different levels in language. Either a single word (lexical ambiguity) or a sequence of words (compositional/structural ambiguity) can have more than one meaning or the context (contextual ambiguity) can establish ambiguity (Löbner, 2015). Regarding lexical ambiguity, we can distinguish two forms. Homonymy is constituted by multiple unrelated meanings of a word like, for instance, “bank” (“financial institute” and “riverside”) and polysemy is constituted by multiple related senses of a word like, for instance, “paper” which can be used to refer to “sheet”, “newspaper”, “publication” and so on (Klein & Murphy, 2001). There is a debate whether the two types of lexical ambiguity differ in the way the meanings/senses are represented. For homonyms it is usually assumed that the distinct meanings are represented separately but for polysemes some studies provide evidence that the senses are separately represented, too (Klein & Murphy, 2001, 2002), whereas others favor a common representation (Beretta, Fiorentino, & Poeppel, 2005; Klepousniotou, 2002; Klepousniotou & Baum, 2007). It is assumed that for some polysemes the senses have a stronger semantic overlap than for others and therefore some behave more like homonyms while others do not (Klepousniotou, Titone, & Romero, 2008). As we use homonyms with clearly distinct meanings for which most authors agree that they are represented separately, we will not focus on polysemy but have a closer look at the process of meaning activation of homonyms in chapter 2.2.

2.2 Processing of lexical ambiguity

It is a matter of interest whether a lexically ambiguous word activates all meanings (exhaustive access) fast and unintentionally or whether the activation depends on the frequency of the meanings (ordered access) or the context (context-dependent access). Compared to studies which show biased interpretation of ambiguity depending on prior processed information (see 2.1) the studies which will be presented in this chapter use the reversed logic: showing a processing bias for some information (the discrete meanings of an ambiguous word) depending on prior processing of ambiguity. In general, the studies investigating the process of meaning activation (or senses for polysemy) of lexical ambiguity, differ in various important aspects: usually they investigate meaning activation for homonymy (e.g., Schvaneveldt, Meyer, & Becker, 1976) but a few studies focus on activation of senses of polysemy (e.g., Williams, 1992). The homonymy studies differ in whether isolated words (e.g., Ishida, 2019) or lexical ambiguity embedded in sentences (e.g. Glucksberg, Kreuz, & Rho, 1986) were used, whether they focussed on balanced (e.g., Holley-Wilcox & Blank, 1980) or unbalanced (e.g., Simpson & Burgess, 1985) meanings, whether they used a visual (e.g., Nievas & Justicia, 2004) or auditory (e.g., Swinney, 1979) presentation mode, and what paradigms they use. Some use behavioral (response time based) paradigms like color naming (e.g., Conrad, 1974), lexical decision (e.g., Frenck-Mestre & Prince, 1997) or pronunciation (e.g., Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982), some apply neurophysiology paradigms like event-related brain potentials by means of electroencephalography (e.g., Klepousniotou, Pike, Steinhauer, & Gracco, 2012) or neuroimaging technics like fMRI (e.g., Bilenko, Grindrod, Myers, & Blumstein, 2009) or eye movement paradigms (e.g., S. A. Duffy, Morris, & Rayner, 1988).

As this diversity shows, there are several ways of investigating the process of meaning activation of lexical ambiguity. One straightforward approach is using semantic priming (for an early review see, Neely, 1991) with the lexical decision task (LDT). In the LDT, participants decide whether a string of letters is a real word or a nonsense word. An early assumption was that participants need to check if an entry in the internal lexicon matches the presented letter string (Rubenstein, Garfield, & Millikan, 1970). A more complex and empirically grounded model that both helps to understand the process of visual word recognition and can deal with the semantic priming effect within the LDT (besides several other findings regarding the lexical decision task) is the dual route cascaded model of visual word recognition (DRC; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). According to this model, semantic priming in the LDT can be explained by the assumption that activation of the semantic representation of the prime (after recognizing the prime including activation of orthographic representation) also pre-activates semantic representation of the target, which influences the activation of the orthographic representation of the target, which leads to faster lexical decision times.

Meyer and Schvaneveldt (1971; Exp. 1) were the first who extended the LDT by simultaneously presenting two semantically related (e.g. bread – butter) or two semantically unrelated (e.g. bread – doctor) pairs of words besides nonword pairs and pairs including one nonword. The participants had to press yes only if the lexical decision for both letter strings was positive; if one or both were a nonword, they had to press no. The important finding was that the lexical decisions for semantically associated words were significantly faster (85ms) compared to unassociated words. It could be assumed that the processing of the first word (the prime) facilitates lexical decision for the other word. For an overview of theories of underlying processes of semantic priming see Wentura and Rothermund (2014).

Schvaneveldt et al. (1976) were the first ones to use the semantic priming effect with a lexical decision task in order to investigate meaning activation of ambiguous words depending on contextual (biased or neutral) information. In their paradigm, participants subsequently saw three letter strings in each trial and should make a lexical decision for each of them. The second word was always an ambiguous word semantically related or unrelated to the first and the last word. They showed facilitation (less time to make a lexical decision) for the last word when the first word primed the same meaning of the ambiguous word as the last word (e.g. SAVE-BANK-MONEY) compared to a control triplet with an unrelated homograph (e.g. FIG-DATE-MONEY). Contrary to this, when the first word primed a meaning which was incompatible with the last word (e.g. RIVER-BANK-MONEY) there was no facilitation compared to the control condition. They interpreted this as evidence for context dependent meaning activation. Triplets with the first word unrelated to the last two words provided a neutral context. The comparison of those triplets in which the last word was either associated with the meaning of the homograph or not, is indicative of whether an ambiguous word primes a meaning at all. The results showed that if the homograph was semantically related to the last word, the lexical decision was faster than if there was no association. It is argued that this facilitation could be due to activation of both meanings within one trial or due to alternating activation of one meaning of the ambiguous word across trials (see also, Holley-Wilcox & Blank, 1980).

Like this study by Schvaneveldt et al. (1976) the majority of studies investigating meaning activation are interested in the influence of contextual information on meaning activation. If the delay between the presentation of the lexical ambiguous word and the words relating to the meaning is taken into account, there are some studies providing evidence that (usually at short intervals between prime and target) even in biased context (when a sentence favors one interpretation of an ambiguous word over the other) both meanings are activated (e.g., Onifer & Swinney, 1981; Rayner, Pacht, & Duffy, 1994; Seidenberg et al., 1982; Swinney, 1979) supporting the exhaustive access account. However, other studies with sentences as context provide evidence that (strong) contextual information has the potential to constrain the initial activation of meanings (e.g., Glucksberg et al., 1986; Simpson, 1981; Simpson & Krueger, 1991), thereby providing evidence for a selective, context-dependent access account. For an early review of this behavioral paradigm in literature dealing with the influence of context for meaning activation see Simpson (1984). A review of studies in which lexical ambiguity arises due to a word having different meanings across languages support an ordered access account (Altarriba & Gianico, 2003). Results from eye tracking paradigms support a hybrid model instead, namely the reordered access model in which context and meaning frequencies of the ambiguous words play a role in the activation of the different meanings over time (for a review see, S. A. Duffy, Kambe, & Rayner, 2001). The model is hybrid in the sense that it posits that all meanings are accessed but the degree of activation depends on both the context and the dominance of meaning. According to this model (similar to the ordered access model) when there is neutral context (not highlighting one meaning) and a balanced homonym is presented, both meanings are activated at the same time.

Results from studies investigating meaning activation in context are not directly comparable to studies investigating meaning activation of isolated words. For instance, different patterns of semantic priming effects using the same ambiguous words in sentences and in isolation were reported by Williams and Colombo (1995). Unfortunately, only a few studies investigated meaning activation of isolated presented lexical homographs without context. In our experiments we use context-free, visually presented homographs with predominantly balanced meanings. That means that the frequencies of words indicating a specific meaning derived from asking participants about their first associations to a homograph do not deviate strongly from each other (Millis & Button, 1989). As this definition of balanced meanings cannot guarantee that both meanings are equally strong associated with the homograph for an individuum, we will now have a closer look at studies providing evidence for or against multiple meaning activation of balanced and unbalanced visually presented isolated (context-free) homographs. While providing this short non-exhaustive review, a discussion of different models of representation would go beyond the scope of this chapter. Therefore, whenever we are referring to meaning activation, we basically mean that a representation of the meaning is activated and hence is more accessible than other meanings and we do not make any further assumptions about the question of how meaning is represented, for instance, by a single unit (Rubenstein et al., 1970) or by a specific pattern of activation of several units (Masson, 1995).

Holley-Wilcox and Blank (1980) investigated whether only one meaning of a homograph with balanced meanings is activated or whether both meanings get activated automatically. One trial in their paradigm consisted of a prime which could be a homograph, an unambiguous word which was either related to the target or a target-unrelated word or a nonword. After conducting a lexical decision for the prime, the target immediately appeared on the screen, for which a fast lexical decision should also be made. The targets were always the meanings of the homographs or nonwords. They found equal facilitation of targets which followed ambiguous related or unambiguous related primes compared to targets following an unrelated prime. This indicates activation of both lexical entries of an ambiguous word on every trial and does not support single (altering) meaning activation on each trial, because the activation of one meaning of a homograph correspond only in 50 percent of trials to the targets and therefore the expected facilitation should be only half the size of the unambiguous related prime condition (for which on every trial the prime was related to the target). The authors interpret their results as strong evidence for multiple meaning activation of balanced lexically ambiguous words.

Simpson (1981; Exp. 1) explored the processing of isolated visually presented homographs with unbalanced meanings. Similar to the study by Holley-Wilcox and Blank (1980), participants' task was to perform a lexical decision for the prime as well as the target. The for the analysis relevant primes in this study were homographs related to the targets and homographs unrelated to the targets. The targets were always the dominant or subordinate meanings of the related homographs. The analysis revealed a relatedness x dominance interaction showing that only dominant related targets were classified faster than the other three possible combinations which needed approximately the same time to be classified. Simpson (1981) interpreted this as evidence against the exhaustive model and in favor of the ordered access account which states that the frequency of associations with an ambiguous word determines the activation of the meanings. However, this paradigm which requires participants to perform a lexical decision task for both prime and target does not allow to control for the stimuli onset asynchrony (SOA) (interval between the onset of the prime and the onset of the target), as it depends on the time to classify the prime. The study by Holley-Wilcox and Blank (1980) reported a mean lexical decision time of 670 ms for the prime which can be regarded as a long SOA.

Simpson and Burgess (1985) investigated the chronology of processing of lexical ambiguity. They used a lexical decision task with ambiguous words (homonyms) as primes. However, in this study a lexical decision should be made only for the targets. Therefore, the presentation time of the prime did not depend on the time needed to perform a lexical decision for the prime. Target stimuli were nonwords, words related to the dominant or subordinate meaning of the ambiguous prime or unrelated words taken out of the stimulus pool from the dominant or subordinate meanings. Furthermore, they varied the SOAs (prime was always replaced by target without interstimulus interval). In Experiment 1, they used SOAs of 16 ms, 100 ms and 300 ms. In Experiment 2 they, used 300 ms, 500 ms and 750 ms. For each SOA they checked for an interaction of dominance with relatedness. Only for the shortest (16 ms) and the longest SOA (750 ms) did they find an interaction showing that facilitation in the lexical decision (mean reaction times for related words compared with unrelated words) was only evident for dominant meanings. For all the other SOAs there was always a significant main effect of relatedness but no interaction with dominance which indicates that for those intervals both the dominant and the subordinate meanings were classified faster in comparison to unrelated words. Simpson and Burgess (1985) interpreted their results as evidence for a multiple-stage process. Only the dominant meaning is activated immediately, then with increasing time the subordinate meaning gets automatically activated, too. After both meanings are activated the subordinate meaning gets inhibited (see also their third Experiment) so that in the end only the dominant meaning remains activated.

Investigating the effect of meaning activation for isolated visually presented homographs (as primes) with unbalanced meanings in native and non-native English speaking samples, Ishida (2019) found facilitation (faster lexical decision compared to unrelated targets) for dominant and subordinate meanings at an SOA of 300 ms for the advanced non-native English speaking sample. In contrast to Simpson and Burgess (1985) the native English speaking sample showed facilitation only for the dominant meaning at an SOA of 300 ms but facilitation for both the dominant and the subordinate meaning at an SOA of 700 ms. Although the time course is shifted these results support the evidence for an ordered access preceding the exhaustive access at a later stage. However, it should be noted that this is the only study we found, providing no evidence for the exhaustive access at medium SOA (about 300 ms) for native speakers.

Several other studies did provide evidence for multiple meaning activation for medium SOAs. For instance, equal facilitation of both meanings of isolated visually presented Spanish homographs in a lexical decision task with an SOA of 250 ms (100 ms prime presentation) for

Spanish university students was found by Nievas and Justicia (2004). Interestingly, they manipulated the balance of meanings using unbalanced and balanced homographs and found no effect of this within factor for this sample. It should be noted that they included children at different ages in the study, too. Evidence for exhaustive meaning activation differed across age groups. Further evidence for exhaustive meaning activation of lexically ambiguous words at an SOA of 300 ms provided a study exclusively with children (8, 10, and 12 years old) using a naming latency paradigm instead of a lexical decision paradigm. (Simpson & Foster, 1986; Exp. 2). At that SOA, both the dominant and the subordinate meaning were facilitated regardless of age of the children. Frenck-Mestre and Prince (1997; Exp. 2) analyzed the lexical decision times for targets words (dominant and subordinate meanings) following a visually presented related homograph or an unrelated word with an SOA of 100 ms and 300 ms for native English speakers, intermediate non-native English speakers and fluent non-native English speakers. For native and fluent non-native speakers, dominance of meanings as well as SOA did not interact with relatedness of the prime. As the facilitation effect (faster responses to related compared to unrelated ambiguous primes) was equally strong for dominant and subordinate primes (for both SOAs) the results support the exhaustive access model for medium SOAs. Only for the intermediate non-native English speakers there was an interaction of dominance of meanings with relatedness of primes. For this group, regardless of SOA, only the dominant meaning was facilitated in comparison to the unrelated prime condition although it was ensured that they knew about both meanings. The authors interpreted this as evidence that the activation of the subordinate meaning was too slow for this group at these SOAs.

Nievas, Justicia, Cañas, and Bajo (2005) investigated the role of associative strength between prime and target words regarding meaning activation of ambiguous words in a Spanish speaking sample. In Experiment 2a, the target words related to the dominant and subordinate meanings had the same associative strength to the homographs which were used as primes. The SOA was 250 ms and the task was to perform a lexical decision task only for the target. They found facilitation for words related to both the dominant and subordinate meaning of a homograph (compared to unrelated homographs as primes) but stronger facilitation for words referring to the dominant meaning. Interestingly, in Experiment 2b they obtained the same pattern of results although the associative strength for words relating to the dominant meanings was twice as strong as the associative strength for the words relating to the subordinate meanings. This between study comparison provides evidence that for short SOAs the factor associative strength can be disregarded when investigating meaning activation of homographs.

In sum, semantic priming studies provide strong evidence that at a medium SOAs (about 300 ms) both meanings of an isolated visually presented homograph are activated regardless of their frequencies. Therefore, in order to investigate the unintentional, fast activation of evaluation of the mental state characterized by having one vs. two meanings in mind, using homographs and non-homographs and compare the evaluative responses to it provide a suitable and parsimonious approach.

2.3 Attitude towards ambiguity

It is necessary to shortly review different conceptualizations of *attitude* before thinking about the implications of these conceptualizations in terms of attitude towards ambiguity. After that, a summary of existing work on direct measures aiming to assess the individual tendency to perceive and respond to ambiguity in a certain way usually labeled as (in)tolerance towards ambiguity is given in 2.3.1. Then we turn to the indirect measurement of attitude towards ambiguity in 2.3.2.

Eagly and Chaiken (1993) provided an umbrella definition of attitude. They define it as “a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor” (p. 1). It could be worthwhile to elaborate the properties of the entity in case of ambiguity. One the one hand, we could make it simple and say that the entity is ambiguity itself (the phenomenon that there are stimuli with more than one meaning). On the other hand, a person could have some degree of liking or disliking of the state of processing an ambiguous stimulus if they have two distinct interpretations in mind. In other words, there may be an attitude towards the phenomenon in general and a specific attitude towards having multiple distinct interpretations or meanings in mind. Since there is this theoretical distinction between these two different entities, it is important to mention which one is the attitude object of interest, the abstract phenomenon or the mental state. It is obvious that subjective ambiguity is not a directly observable entity like other common attitude objects e. g. persons, products, ethnic groups, political parties et cetera. Therefore, ambiguity could be considered as an abstract attitude object in the sense that it has no single physical appearance and resembles other highly abstract and not directly observable attitude objects such as religion or science. However, subjective ambiguity differs from these abstract attitude objects such that the individual content of the distinct representations is irrelevant because subjective ambiguity is defined by the mental state of coactivation of distinct representations. Therefore, regardless of the content these representations, the attitude object is identical across subjects and could be considered as a very concrete attitude object.

In direct measures the content of the item predominantly implies that evaluative responses for the abstract phenomenon are assessed. As we will see, some indirect measures may be better suited to further investigate the evaluative response only to the ambiguousness of stimuli and, inevitably connected with this, the assumed triggered mental state of having two (or more) interpretations in mind. As mentioned above, the activation of at least two different representations triggered by a stimulus is necessary to say that this stimulus is ambiguous for a person. A problem with investigating the evaluative response only to the ambiguousness of stimuli is that ambiguity never occurs by itself but always needs a medium like a word, a phrase or a social action. However, this medium and its separate meanings/interpretations could carry some valence too and hence be attitude objects as well. Considering the word “bank”: On the one hand, it refers to a financial institute and on the other hand, it means a riverside. If we try to express the “degree of favor or disfavor” only for the state of having the mentioned two different interpretation in mind, we face the problem that the valence of the two meanings are coming to mind too. This example shows that evaluation of ambiguity is not trivial.

But what is evaluation? Some researchers define evaluation as the thoughts, beliefs, and judgments about an attitude object and clearly separate affect (the emotional responses triggered by an attitude object) from evaluation (Breckler & Wiggins, 1989), although affect, behavior, and cognition are discussed as distinct components of attitudes (Bagozzi, 1978; Breckler, 1984; Rosenberg, Hovland, McGuire, Abelson, & Brehm, 1960). According to De Houwer (2009) evaluation (the process of identifying the liking of an object) is not directly observable and the study of evaluation can only be done by investigating evaluative responses driven by an object. We can infer an evaluative process if an evaluative responding took place (De Houwer, 2009). Under this conceptualization, affect is part of evaluation because the responses can have different manifestations like for example physiology, behavior, selecting a point on a scale or the performance in indirect tasks. The focus on evaluation which determines different evaluative responses makes it possible to look at automatic (fast, unintentional and efficient) evaluations of ambiguity assessed with indirect measures, in addition to direct verbal evaluation, in order to infer the attitude towards ambiguity. Especially as ambiguity is such an abstract phenomenon and because the valence of ambiguity is confounded with the valence of the meanings, observing different evaluative responses allows for a deeper insight into the psychological tendency expressed by evaluating ambiguity with some degree of favor or disfavor.

One question in the ambiguity intolerance literature is, whether there really is one general attitude towards ambiguity or several attitudes towards several domains in which ambiguity occurs (see 2.3.1). Besides this question, there is an even more fundamental debate about attitudes. The research differs in the way it assumes that attitudes are stored in memory or constructed in a specific situation (Bohner & Dickel, 2011). On the one hand, attitudes could be considered as stored associations between the representation of an object and the representation of valence in a sense that evaluative summaries of different aspects about an object are associated with the object (Fazio, 2001, 2007; Fazio et al., 1986). On the other hand, Schwarz (2007) treats attitudes as evaluative judgments constructed on the spot. He argues that they are formed whenever the situation demands for an evaluation instead of being there all the time like a disposition. The constructionist conceptualization is built on the assumption that only context sensitive evaluation is suited for adaptive behavior. Reviewing the literature, he showed that the constructionist conceptualization can account for a bulk of empirical finings. From this point of view, true attitude does not exist because: “With variation in context, multiple evaluations of an attitude object may be evoked, but none of those evaluations is more true than any other” (p. 468) (J. P. Mitchell, Nosek, & Banaji, 2003). Without arguing for or against one account it is worthwhile to consider these two conceptualizations of attitudes when thinking about attitude towards ambiguity. Especially when it comes to generating an item pool for constructing a direct measure of attitude towards ambiguity, this distinction between constructed or stored attitudes can influence to which degree one includes contextual variation in order to capture dissimilar evaluative judgments across different contexts. Considering systematic contextual variations early in scale development (e.g. at the stage of item generation and before exploratory factor analysis) allows for exploring whether the evaluative judgments are the same for all domains of ambiguity or differ across domains. This information then allows to speculate (from an attitude as latent construct point of view) about the existence of one attitude towards ambiguity or several attitudes towards several domains of ambiguity. The domain (in)sensitivity of attitude towards ambiguity is discussed in literature and we focus on this point in 2.3.1.

With the appearance of new indirect response time-based measures of attitudes (see 2.3.2), considerations of different underlying processes between direct and indirect measures result in new models of attitudes and attitude change. We do not adopt the usual labels “implicit/explicit” measures because this distinction usually implies the assumption of different functional properties (like controllability, awareness, intentionality etc.) which are determined by the conditions under which the assessment takes place (De Houwer, 2008). Corneille and Hütter (2020) reviewed the attitude research using the “implicit” terminology and noted that this does not only imply assumptions about the automaticity by some authors but assumptions about the mental processes (e.g. associative) by others or simply refer to the procedure of measurement (indirect) or combinations of conceptualizations making the term “implicit” ambiguous, in the core sense of ambiguity. The distinction “direct/indirect” is based on an objective difference: does the participant directly self-assess the construct of interest or is that construct indirectly inferred from other responses (e.g. response times) (De Houwer, 2008).

One prominent model of attitude and attitude change falling in the class of dual process models is the associative-propositional evaluation (APE) model (Gawronski & Bodenhausen, 2006, 2011, 2014). The APE model claims that evaluative responses can result from two qualitatively different processes. Automatically activated associations when processing an attitude object result in automatic affective reactions. Associative evaluations therefore form the basis for attitudes assessed via indirect measures. The process of propositional reasoning serves to validate the propositional information relevant for the evaluative judgment. The propositional information derives from other propositions or from transformed activated associations. The result of the validation process usually builds the basis for attitudes assessed via direct measures of attitudes. The associative and propositional process can influence each other, and the model allows for predictions when attitudes assessed via direct and indirect measures correspond with each other or deviate from each other (see Gawronski & Bodenhausen, 2006, 2011).

However, Albarracín, Hart, and McCulloch (2006) argue that it is not necessary to postulate two qualitatively different processes to account for the findings. Instead, they argue that the propositions could be ordered associations. On a more general level, J. W. Sherman, Krieglmeyer, and Calanchini (2014) criticized the usual practice of dual process accounts, namely to infer the underlying operating principles (associative vs. propositional) from the operating conditions (automatic vs. controlled) of measures. One main problem of this reasoning is that there is no process purity (automatic or controlled) of any indirect or direct measure (J. W. Sherman, 2008) see 2.3.2.

Despite the lack of clarity concerning the underlying operating principles in attitude formation and change, the use of direct and indirect measures of attitude towards ambiguity allows for different conclusions without making assumptions about the underlying operating principles. On the one hand, some indirect measures (e.g. the EP paradigm) are suited to investigate the unintentional fast activation of evaluations of stimuli under special conditions. We use the EP paradigm to investigate whether ambiguity or clearness activate evaluations unintentionally and fast. To our knowledge, this topic has never been addressed on an empirical basis using indirect measures. On the other hand, as direct measures are more reliable than indirect measures, they are more appropriate to investigate the structure of attitude towards ambiguity in the sense that they allow to investigate the domain specificity and the relation to other related attitudes. Researchers have dealt with these issues since the middle of the last century but failed to apply a narrow definition of ambiguity.

Although it is not in the focus of our research, we will investigate the relation of direct and indirect measures of attitude towards ambiguity expecting a low correlation as correlations of these types of attitude measures are usually small (see, Cameron, Brown-Iannuzzi, & Payne, 2012; Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005). However, it should be noted that the strength of association between these types of measures depends on the topic. Calculating the correlation between latent variables representing the explicit and implicit attitude (assessed via the IAT) across 57 topics, Nosek and Smyth (2007) report correlations ranging from .01 to .79. To our knowledge, there is no reported correlation using common response time based indirect measures with explicit measures for ambiguity. There is only evidence that there is no correlation of direct (questionnaire based) and other indirect (via text analytic methods) measures of tolerance of ambiguity (Leichsenring, Steuernagel, Steuernagel, & Meyer, 2007).

2.3.1 Direct measures of attitude towards ambiguity

Personality trait or attitude? As we will see, the tendency to perceive and respond to ambiguity is most often seen as a personality variable (see Furnham & Marks, 2013; Lauriola, Foschi, Mosca, & Weller, 2016). Eagly and Chaiken (1993) argue that attitudes are distinguishable from personality traits because on the one hand the latter covers much more diverse stimuli that potentially are responsible for a response, and on the other hand these responses are broader and more diverse than pure *evaluative* responses triggered by an attitude object. Applying this discriminative rule to ambiguity, one could argue that nearly all stimuli or situations have the potential to be ambiguous which on the surface suggests that the tendency to react to ambiguity is considered a personality trait. However, considering that ambiguity is constituted by the state of having multiple distinct interpretations or meanings in mind, one could argue that this is indeed a very specific “stimulus”. That the tendency to react to ambiguity can best be conceptualized as an attitude is further backed up by the assumed responses to ambiguity. Nearly all responses postulated in definitions of (in)tolerance of ambiguity or inferred from items of the scales (see below), irrespectively of cognitive, affective or behavioral content, imply an underlying evaluation and can thus be regarded as evaluative responses. Therefore, we argue that the evaluative tendency to ambiguity suggests that it should be considered an attitude and not as a personality variable. It should be noted that the special view of treating the ambiguity as an attitude object only makes sense when applying a narrow definition of ambiguity. The way the concept is treated in most direct measures (as an umbrella term for dissimilar types of situations) allows to consider the tendency to react to these situations as a personality trait.

Overview of the concept and measures of (in)tolerance of ambiguity. Frenkel-Brunswik (1948, 1949) introduced the concept (in)tolerance of ambiguity (TA) to the literature. She treated the concept as a personality trait referring to emotional and perceptual aspects. While analyzing data from 120 school children extremely high or low in ethnic prejudice, she found that some children directly reported only positive aspects about their parents although indirect measures provided evidence that they indeed did have negative associations with their parents, too. Other children did recognize and report positive and negative aspects of their parents. She argues that the personality trait TA accounts for the finding that some have the ability to recognize and report this “coexistence of positive and negative features in the same object” (p. 115) (Frenkel-Brunswik, 1949). In that sense, TA should determine the amount of ambivalence, that is the simultaneous activation of opposite (positive and negative) evaluations (Berger, Hütter, & Corneille, 2019). She investigated if TA is only linked to that emotional domain or if it also refers to perceptual phenomena and more cognitive issues. Reporting experiments dealing with biased memory and perception phenomena, she concluded that the prejudiced children indeed showed a tendency to reduce the amount of ambiguity in these more cognitive tasks too. She even conceptualized the trait as being bipolar, ranging from high tolerance towards ambiguity which means a high “ability to recognition such coexistences” (p.115) to high intolerance towards ambiguity which is defined by “a tendency to resort to black-white solutions, to arrive at premature closure as to valuative aspects, often at neglect of reality, and to seek for unqualified and unambiguous over-all acceptance and rejection of other people” (p. 115). Frenkel-Brunswik (1949) was also concerned with the question if TA is a “formal characteristic of the organism independent of context” (p. 126). This question of context dependency is prevalent in literature as we will see. Since the first mention of the concept TA by Frenkel-Brunswik (1948) and especially since her contributions dealing with the status of TA in relation to other personality constructs and with the integration of the concept into the cognitive architecture of personality in “The Authoritarian personality” by Adorno, Frenkel-Brunswik, Levinson, and Sanford (1950), the concept has aroused great interest in several domains of psychology and economics (see for a review of TA, Furnham, 1994; Furnham & Marks, 2013; Furnham & Ribchester, 1995; Herman, Stevens, Bird, Mendenhall, & Oddou, 2010).

In the following section we will focus on the development of direct measures of TA including consideration of the dimensional structure of the construct. Although most authors treat the concept as unidimensional, factor-analytic studies with multiple direct measures showed that it is better represented by multiple dimensions (Furnham, 1994; Lauriola et al., 2016). As we will see, there is a huge diversity in what is meant by TA. Unfortunately, many authors did not define TA, but still have engaged in developing direct measures for it. From introduction of the concept until now, there was only marginal theoretical development of the construct TA (Furnham & Marks, 2013). One of the first attempts to define TA was given by English and English (1958). They defined TA as "…willingness to accept a state of affairs capable of alternate interpretations, or of alternate outcomes: e.g., feeling comfortable (or at least not feeling uncomfortable) when faced by a complex social issue in which opposed principles are intermingled" (p. 24). In contrast to the definition of TA by Frenkel-Brunswik (1949), they focused on the emotional response to ambiguity. Since then, this would remain an important component of definitions of TA. Kenny and Ginsberg (1958) conducted a study with 12 different measures of TA in order to investigate the assumption of an underlying unitary construct empirically. Among these measures of the construct they found nearly no intercorrelations. They interpreted the result as indicating that there is no evidence of a common factor of TA underlying these measures. However, it should be noted that the measures are potentially highly unreliable. For instance, the only used direct measure, the Walk A Scale (first mentioned in an unpulished manuscript but for a reproduction see, O'Connor, 1952) is highly unreliable in terms of internal consistency (K-R 20 of .08, Ehrlich, 1965). On the other hand, one could argue that the measures (e. g. number of questions asked during the experiment, number of fluctuations in reversible figures) are not valid measures in a sense that they are not assessing what they are supposed to assess (the construct TA). Probably because of lack of published and validated scales of TA, researchers interested in investigating relations to other traits (like ethnic prejudice) constructed their own direct measures without reporting much about the process of item generation, selection and evidence for reliability and validity (see, Martin & Westie, 1959).

The first published and commonly used scale assessing TA stems from Budner (1962). A bipolar conceptualization with intolerance and tolerance as endpoints of a unidimensional scale is assumed. Intolerance of ambiguity is defined as "'the tendency to perceive (i.e. interpret) ambiguous situations as sources of threat', tolerance of ambiguity as 'the tendency to perceive ambiguous situations as desirable'" (p. 29) (Budner, 1962). According to him, situations are ambiguous when they are completely new, too complex or have contradictory elements and hence are unsolvable. If one of these situations elicits feelings of anxiety, avoidance or destructive behavior or a general denial of the situation, he states, the degree of intolerance of ambiguity is the underlying personality trait responsible for intraindividual variations of these reactions. The 16 Items are designed to tap into one type of response and refer to one type of situation described above. He reports reliabilities (Cronbach's alpha) for 13 samples (N between 33 and 86) ranging from .39 to .62 (mean alpha is .49) and a test-retest correlation of .85 for one sample (N = 15). Intercorrelations to three other direct measures of TA ranged from .36 to .54 (highest with Walk's A Scale) and are interpreted as assessing the same construct and therefore as evidence for validity. According to Budner (1962), evaluation is central to TA and he points out that the diverse responses to diverse situations should be seen as manifestations of that trait. It must be noted that with the given definition, the concept is assessed in a broad sense. For instance, the item no. 9 "I would like to live in a foreign country for a while" could also be an item for measuring openness to new experience, which is one of the five dimension of the personality inventory NEO-PI-R (Ostendorf & Angleitner, 2004). Intuitively, it seems doubtful that all 16 items represent the same construct and hence load on a single factor. Empirical evidence against a single factor model comes from Furnham (1994) who conducted an exploratory factor analysis and found a four factor solution. Nevertheless, confirmatory factor analysis provided evidence against Furnham's (1994) proposed four factor model (Benjamin Jr, Riggio, & Mayes, 1996). However, Budner's (1962) assumed one-dimensionality model showed an unacceptable fit to different data sets, too (Benjamin Jr et al., 1996; Bors, Gruman, & Shukla, 2010).

Primarily in order to investigate the relation of TA with need for cognition (defined as the need to know or understand), S. T. Rydell and Rosen (1966) developed another direct scale comprised of 16 items with a true-false response opinion. They report test-retest reliabilities of .71 (for a one-month interval; N = 41) and an *r* of .57 (for a two-month interval; N = 105). The development of the scale was done explicitly without recourse to definitions and "on a priori basis" (p.150). Evidence for the validity is provided by a study by S. T. Rydell (1966). She could show that a low (vs. high) TA group based on that scale had higher extremity ratings of contradictory adjective-noun pairs (i.e. beautiful abortion) indicating that they those people tended to disambiguate in a sense that they focused more on one component. The high TA group, however, seemed to integrate these two conflicting meanings instead. The majority of items does not fit a narrow definition of ambiguity and may assess more how much a person tolerates uncertainty in different situations. Although intolerance of uncertainty and TA indeed have similarities like an evaluative tendency (usually threat) to respond to specific situations, the two concepts can be theoretically distinguished by regarding which information elicits the response: a present information with multiple interpretations in case of ambiguity or an unpredictable outcome in the future (Grenier, Barrette, & Ladouceur, 2005). Importantly, S. T. Rydell and Rosen (1966) mentioned a new aspect of TA when pointing out the difference between tolerance of the phenomenon in situations where it is essential and an integral part of the situation (like in abstract art) and tolerance of the phenomenon in order to achieve another goal (like in science where ambiguities arise in the process of gaining knowledge). We will pick up this distinction while constructing a new scale where the ambiguity in art is considered a special case because it is often an integral part of art. Only one item of their scale deals with ambiguity in art ("Vague and impressionistic pictures really have little appeal for me").

Mac Donald (1970) added four items to the 16 items from the S. T. Rydell and Rosen (1966) scale in order to improve psychometric properties. Referring to existing definitions he considers "that persons having high tolerance of ambiguity (a) seek out ambiguity, (b) enjoy ambiguity, and (c) excel in the performance of ambiguous tasks" (p. 791). For the original 16 items he reported a split half correlation (Spearman-Brown corrected) of .64 and for his extended (AT-20) version of the scale he reported an r of .86 with the same reliability estimate (N = 74). For a second, bigger sample (N = 789), he reported an internal consistency of .63 (K-R-20 formula) for the AT-20 scale. The AT-20 scale outperforms the other existing scales with respect to internal consistency but especially taken the relative high number of items into account the reported consistency estimation speaks for a rather low average item correlation (see Cortina, 1993). Furthermore, he reported a test-retest reliability of .63 (interval of six months). With the purpose of providing support for predictive validity, he reported a correlation (r = .33) of the AT-20 scale with the performance of solving anagrams. The component enjoyment of ambiguity which is part of his view on properties of TA, was not validated. Instead, he focused on the construct validity. Correlations to related constructs are reported. TA correlated significantly with dogmatism (r = -.42) and with rigidity (r = -.41) which he interpreted as evidence that the three concepts have a common dimension although he previously argued that "the two constructs [TA and rigidity] are theoretically and empirically separate" (p. 791). A significant correlation of TA and church attendance (r = -.24) is supposed to further support validity. Despite of clear improvements on measuring TA, continuing effort was made in order to further improve the psychometric properties and provide broader validity evidence.

A next step in the direction of developing a reliable and valid instrument for assessing the TA was made by Norton (1975). Reviewing the usage of the term ambiguous in the literature, he came to this working definition: "Intolerance of ambiguity is a tendency to perceive or interpret information marked by vague, incomplete, fragmented, multiple, probable, unstructured, uncertain, inconsistent, contrary, contradictory, or unclear meanings as actual or potential sources of psychological discomfort or threat” (p. 608). Compared to the three situation types which Budner (1962) postulated to be ambiguous, this understanding of ambiguous is even more extensive, trying to cover all existing applications of meanings of the term ambiguous. He focused on the emotional-affective consequences (threat and discomfort) of these situations for a person with high intolerance of ambiguity. Therefore, the items were constructed to reflect the four proposed consequences of Budner (1962), namely avoidance or destructive behavior, feelings of anxiety or a general denial of the situation. Although no factorial structure is presumed, the items reflect eight different domains (philosophy, interpersonal communication, public image, job-related, problem-solving, social, habit, art forms). The scale development followed the goal of high internal consistency using the K-R-20 formula. The seventh version of the scale (MAT-50) with 52 items reached a high internal consistency of .88 (N = 208). Norton (1975) provided evidence for content validity, construct validity (significant correlations to other related constructs like rigidity and with other TA scales), predictive validity (e.g. the scale predicts (1) the agreement of taking part in another not further specified study, (2) evaluations of poems with different degrees of ambiguity and (3) Amount of ambiguous speech in a communication context). On the one hand the MAT-50 could be seen as a further improvement in assessing TA, but on the other hand it is not the most economic instrument with its 52 items. Interestingly, a distinction between objective and subjective ambiguity is proposed. Objective ambiguity refers to stimuli which are inherently ambiguous due to their structure, although that does not mean that the ambiguity is seen. Subjective ambiguity depends on the interpretation of a (not necessarily objective ambiguous) stimulus.

Kischkel (1984) developed a more economic TA scale comprising of 14 items derived from the MAT-50 scale (Norton, 1975) and the AT-20 scale (Mac Donald, 1970). Internal consistency of .76 (K-R-20) and evidence for convergent and discriminant validity is given. Because of dominance of job-related ambiguity, the scale does not seem to be a valid instrument for general TA. Other very specific TA scales assessing TA within the field of second language use were constructed by Ely (1989).

Another attempt to develop a short but reliable instrument for assessing TA was made by Geller, Tambor, Chase, and Holtzman (1993). Performing item and exploratory factor analyses in subsequent samples on (in some cases slightly modified) items from existing scales (Budner, 1962; Mac Donald, 1970; Norton, 1975) they came up with a 7 item scale with an overall Cronbach’s alpha of .75 (N = 1420), which is acceptable considering the low item number. Most interestingly, although not intended, they found a two-factorial structure and interpreted the first factor as “desire for certainty” including 4 items and the second factor as “willingness to admit discomfort with ambiguity” (p. 995). This is the first empirical evidence that these two traits may have separate underlying latent variables. How they are connected is not reported by the authors. It could be assumed that there is a relation because on the one hand certainty and ambiguity are natural opposites, and on the other hand the axis was rotated obliquely. The authors used the terms uncertainty and ambiguity interchangeably which can be criticized (see, Grenier et al., 2005).

McLain (1993) constructed a 22-item scale with high internal consistency (Cronbach’s alpha of .86). Unlike other authors he exclusively used newly designed items free from associations to specific domains or concrete situations. The items cover the intended construct on an abstract level (e. g. item 1 “I don’t tolerate ambiguous situations well”). In order to achieve high content validity of items he clearly defined separately what he considers as ambiguity and what he means by tolerant. Ambiguity is the “perceived insufficiency of information regarding a particular stimulus or context” (p.183) and “tolerance extend[s] along a continuum from rejection to attraction” (p. 184). According to him, ambiguity can result of perceived unfamiliarity, complexity, unpredictability, novelty or if a stimulus elects multiple interpretations. The items cover these determinants of ambiguity (e. g. item 21 “I enjoy an occasional surprise”). Convergent validity is provided via significant correlations to existing scales of TA (ranging from .37 to .58). As the authors before did, too, correlations to theoretically related constructs like dogmatism ($r = -.34$) are regarded as evidence for validity. Based on factor loadings he proposed a single dimension underlying the trait (factor loadings between .18 to .84). How many factors should be retained applying the Kaiser-Guttman criterion or regarding to the Scree test is not reported. The big range of loadings (from .18 to .84) and a substantial number of items only loading relatively weak (e. g. 6 items below .40 and 12 items below .50 on the first factor) could be taken as evidence that the data may be better represented by multiple factors. Additional confirmatory factor analyses in a separate study would have provided a better understanding of the factorial structure of that measure. In a revised version of this scale (see below) this was done. Giving no additional instruction to participants of what ambiguity is and asking them about the degree of endorsement of statements like “I try to avoid situations which are ambiguous.” (Item 6) is problematic because people may have completely different conceptualizations of ambiguity (for an overview of different meanings of 'ambiguous' see Norton, 1975).

Webster and Kruglanski (1994a) developed a questionnaire (see, Webster & Kruglanski, 1994b) to assess the need for cognitive closure. This trait covers five aspects with one of them being discomfort with ambiguity. Reported internal consistencies (based on Cronbach's Alpha) for the nine items representing this facet of the scale are .67 (N = 281) and .80 (N = 172). Two of the nine items (item 30 "I dislike it when a person's statement could mean many different things." and item 36 "I feel uncomfortable when someone's meaning or intention is unclear to me.") are compatible with an understanding of ambiguity as multiple interpretations/meanings of a stimuli. Other items (e. g. item 29 "I like to know what people are thinking all the time") do not seem to assess either discomfort or an evaluative response to ambiguity at all. In line with other TA measures, convergent validity is provided via correlations with another measure of TA (r =. 40) and with other related concepts like dogmatism (r = .32). Although one could argue that TA is a broad trait based on the literature, Webster and Kruglanski (1994a) treat it solely as one manifestation besides others of an even broader domain unspecific unidimensional disposition. From the perspective of TA literature, the other facets (preference for order, preference for predictability, decisiveness and closed-mindedness) of the trait need for closure could be regarded as manifestations of TA, too. Many items of these facets are similar to items from TA measures. This illustrates that there is no agreement with respect to localization of the trait TA in the hierarchy of personality traits as superordinate or subordinate construct.

Picking up the idea of different domains of TA proposed by Norton (1975), Reis (1996) made great efforts (6 samples with a total N = 1665) to discover and validate discrete dimensions representing different domains of TA using factor analytic procedures. He refrained from providing a definition of ambiguity or TA. Instead he started the scale construction by asking participants about ambiguous situations and designed items that reflect these situations. After several factor and item analyses, the total inventory consisted of 40 items divided in 5 scales reflecting TA in: social conflicts (6 items), view of parents (11 items), gender stereotypes (9 items), new experiences (6 items) and unsolvable problems (6 items). The reported indicators of reliability, internal consistency (Cronbach's alpha of .88; N = 169) and stability (test-retest reliability of .97; N = 169), are higher than those of other TA scales. Evidence for validity comes from correlations with other TA measures. In contrast to the context independent approach from McLain (1993), the items of these scales are highly context dependent (e. g. item 2. "Frauen sollten sich beim Tanzen vom Mann führen lassen"; translation: "Women should be led by the man while dancing"). In most cases the content validity of items, especially referring to narrow definition to ambiguity, is doubtful (e. g. item 24. “Meine Eltern haben mich zu wenig geliebt”; translation: “My parents loved me insufficiently”).

Durrheim and Foster (1997) deal with the question if TA is a generalized personality trait or varies across different domains. Their understanding of TA goes back to Frenkel-Brunswik’s (1949) finding that there is a variation in reporting positive and negative aspects of one’s own parents (see above). Even though she postulated that ambivalence is only one among several other perceptual/cognitive manifestations for which TA could be the underlying trait, Durrheim and Foster (1997) focused on that specific understanding of TA. For measuring ambivalence they used an objective ambivalence formula (see, Scott, 1966) which calculates the amount of ambivalence out of two unipolar scales separately assessing the liking and disliking of an object. Doing that for 20 objects (e.g. Jesus, Doctors, Parents) and conducting an exploratory factor analysis with these values revealing four factors, they concluded that TA is a content specific construct. In our view, focusing only on the amount of ambivalence participants have over several domains of social stimuli and the result of partially unrelatedness of ambivalence is not suited to investigate the domain specificity vs. generality of TA. Furthermore, one could argue that they mix up the ability to have ambivalence with the tolerance of ambiguity. Unfortunately, their results are constantly being interpreted in terms of evidence for inter-domain differences in TA.

In 2009, (McLain) modified his previous scale (McLain, 1993). Like before, he defined tolerance as bipolar with attraction and aversion as endpoints of one dimension. To the three ambiguous stimulus types (unfamiliarity, complexity and insolubility) proposed by Budner (1962) he added uncertainty. The 13 context-free items derived from his first scale tap into these four types of ambiguous situations or refer to not further specified ambiguous situations (e. g. item 7. “I am tolerant of ambiguous situations”). All types are regarded as involving ambiguity as they all have in common that there is temporarily insufficient information for comprehension of the situation or to “identify its possible future states” (p. 977). The last part of his definition of ambiguity refers to uncertainty, which can be seen as conceptually separate from ambiguity (Grenier et al., 2005). The first part of the definition allows the different situations mentioned to be subsumed as ambiguity. The scale shows a good internal consistency indicated by Cronbach’s alpha of .83. ($N = 542$). A confirmatory factor analysis provided evidence for a one-dimensional model but only if error terms were allowed to covary. In order to provide construct validity he conducted a further study ($N = 121$) and reported significant correlations ($r = .41$) to the AT-20 scale (Mac Donald, 1970) and to sensation seeking ($r = .27$) which is considered a related construct. Interestingly, there is no substantial correlation ($r =$

.09) to the theoretically most related scale by Budner (1962). The low reliability of the Budner scale should be considered. In a third validation study with 207 persons working in high risk jobs, he reported significant positive correlations to risk orientation, perceived risk, perceived uncertainty and a significant negative correlation to stress. He considers the data to be supporting the view of TA as a general personality trait.

The most recently published scale of tolerance for ambiguity (TAS, Herman et al., 2010) used the 16 item scale from (Budner, 1962) as basis plus 5 new items with high cross-cultural relevance. Data for the 21 items was collected in a large sample (N = 2351). In a first step the authors reduced the Budner scale to 12 items based on low item-total correlation. Then they added the 5 new items and excluded 5 more Buder items based on low correlation with the total score. The 12 resulting items (including 7 Buder items) were analyzed by principal component analysis with oblimin rotation. They extracted 4 interpretable factors and labelled them as: “(1) valuing diverse others, (2) change, (3) challenging perspectives, and (4) unfamiliarity” (p. 62). The fist factor contains three new items (e.g. item 1 “I avoid settings where people don’t share my values”). Although they criticized the broad spectrum of conceptualizations, they did not clarify their understanding of the concept and in which sense the new items represent an ambiguous situation. They referred to the definition of McLain (1993) in which stimuli with multiple interpretations (which we considered as the core of ambiguity) are an aspect of ambiguity. Unfortunately, none of their items refers to multiple interpretations. Furthermore, their goal to develop a psychometrically superior scale has failed, considering the internal consistency indicated by Cronbach’s alpha of .73 of the total score and lower than or equal to.58 for the four proposed dimensions.

Summary of direct measures of TA. All scales have in common that diverse types of situations are regarded as ambiguous, and as far as we know, there is no direct measure applying a narrow definition of ambiguity in which ambiguity is characterized only by multiple interpretations. So, the conceptualization of ambiguity in direct measures is ambiguous because different authors have different and multiple interpretations of the construct. This extensive conceptualization of TA reflected in items of the scales could lead to the correlations found with many other personality traits or behaviors (for an overview see Furnham & Marks, 2013). Furthermore, most scales lack an acceptable level of reliability and those which are reliable consist of many items. An exception is the scale from McLain (2009) which is economic, reliable and, additionally, seems to have only one underlying latent variable. But the broad understanding of ambiguity underlying this scale should be regarded as problematic, too. Especially so when the scale is used to measure the attitude towards ambiguity in a sense that ambiguity is characterized by having multiple interpretations or meanings in mind, or more generally, towards the phenomenon that there are stimuli or situations with more than one interpretation or meaning. Moreover, the question if there are intraindividual differences in the attitude towards ambiguity in different domains remains unsolved. Referring to different domains, in contrast to other authors, we do not mean different aspects of ambiguity (e. g. unfamiliarity, complexity) but that the same phenomenon (multiple meanings/interpretations) occurs in different domains, for example in a job context or in art. Therefore, the development of an economic, reliable scale with high content validity and potential factor-analytically justified subscales representing the attitude towards the same construct (ambiguity) but in different domains is desirable. In chapter 3.2 we report the development of such a direct measure.

A key element of most definitions or the content of the items implies that if a person is highly intolerant towards ambiguity, this person usually perceives ambiguous stimuli as negative or, more specifically as “potential sources of psychological discomfort or threat” (p. 608) (Norton, 1975). On the other hand, persons with high tolerance of ambiguity perceive ambiguous stimuli as positive or, more specifically “enjoy” ambiguity (MacDonald, 1970). Interestingly, this core assumption that there are interindividual differences in the “affective” or evaluative response when having multiple meanings in mind, depending on the interindividual differences in the attitude towards ambiguity is not empirically validated.

We investigated this question in two ways: On the one hand, while developing a new direct measure of attitude towards ambiguity (chapter 3.2), we investigated this question with direct measures of evaluation of ambiguous and unambiguous stimuli/situations and direct reported affect. On the other hand, with an indirect measure (the EP paradigm), we test the hypothesis that there are interindividual differences in automatic activation of evaluation of ambiguity in line with the indirectly assessed attitude towards ambiguity (assessed via the IAT) an in line with a previously manipulated attitude. In the next chapter we discussed why this approach (and why these indirect measures) are best suited for our question.

2.3.2 Indirect measures of attitude towards ambiguity

While facing the limits of direct measures, especially in the case of assessing evaluative responses to ambiguity, Budner (1962, p. 32) offers the view that “the most accurate estimate of an individual's degree of tolerance-intolerance of ambiguity would be obtained by extended observation of spontaneous responses to "real-life" ambiguous situations”. We claim that this procedure does not only allow to estimate the trait but also to validate the assumption that there are indeed spontaneous evaluative responses to ambiguity.

One precursor of the new indirect measures was Zajonc (1980) who stated that there is a separate mental system for affective information enabling a fast, higher cognition preceding, affective reaction to stimuli. He built his conclusion primarily on the observation that previously seen stimuli were liked more than unseen stimuli despite participants not being able to distinguish between old and new stimuli. Murphy and Zajonc (1993) gave further evidence supporting the affective primacy hypothesis. They showed that during subliminal stimulus presentation only affective (vs. cognitive) primes lead to evaluative change of novel stimuli, whereas during longer prime presentation only cognitive primes influenced judgments. Chen and Bargh (1999) conducted experiments that showed that objects automatically produced motoric predispositions (pull or push al lever). Thereby, they gave evidence that automatic evaluation can influence our approach and avoidance behavior. Duckworth, Bargh, Garcia, and Chaiken (2002) summed up the properties of evaluative responding: it “can be immediate, unintentional, implicit (i.e., occurring without awareness), stimulus based, and linked directly to approach and avoidance behavioral tendencies” (p. 513). Their experiments provide evidence that even new objects (without preexisting object evaluation associations) can trigger an immediate and unintentional evaluation by using them in an evaluative priming paradigm. The generality of automatic evaluation (see Bargh, Chaiken, Govender, & Pratto, 1992) in terms of Zajonc (1980) labeled as automatic affective reaction could be seen as the basis for the new indirect measures usually based on response times.

Indirect measures of attitudes have been designed to overcome strategically, deliberate influences that could bias the assessment of the “true attitude” which is especially a problem in social sensitive attitudes like prejudices (Fazio et al., 1995) on the one hand and introspective limits about processes of judgements and behavior (Nisbett & Wilson, 1977) on the other hand. These measures have in common that they are based on the assumption that evaluations of an object can be activated automatically (Fazio, 2001; Fazio et al., 1986; Hermans, Houwer, & Eelen, 1994; Herring et al., 2013). This *activation* of evaluation is assumed to occur automatically in a sense that it is unintentional, without awareness, efficient and uncontrollable without the need for conscious or deliberate processing of the object (see De Houwer, 2008, for the importance of investigating the functional properties of indirect measures). Thus the process refers to a sort of preconscious automaticity (Bargh, 1994). Of course, some of these properties of automaticity are now known to depend on specific circumstances (see below). Therefore, when we refer to the term automatic, we only refer to the features unintentionality and fastness.

Unintentionally means that the process takes place even when there is no goal to start the process (Moors & De Houwer, 2006). Several indirect measures emerged that are built on the assumption of automatic activation of evaluations. The established ones are the evaluative priming (EP) paradigm (Fazio et al., 1986), the emotional stroop task (Pratto & John, 1991), the implicit association test (IAT; Greenwald et al., 1998) and the single category implicit association test (SC-IAT; Karpinski & Steinman, 2006), the affective simon task (De Houwer & Eelen, 1998) and the extrinsic affective simon task (EAST; De Houwer, 2003a), the Go/No-Go association task (Nosek & Banaji, 2001) and the affect misattribution procedure (AMP; Payne, Cheng, Govorun, & Stewart, 2005).

The usual finding is that indirect measures are not highly correlated with each other (see Nosek, Greenwald, & Banaji, 2007). Implying that either the measures are unreliable and/or that they are assessing different constructs or are based on different cognitive processes. Already Klauer and Musch (2002) claimed that indirect measures “may be differently responsive to situational, attentional, and even motivational factors, and they may differ in the degree to which they can be strategically controlled” (p. 813). Meanwhile it is well accepted that no measurement is process-pure (J. W. Sherman, 2008). In addition to automatic processes indirect measures reflect another process too (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005; Klauer, Voss, Schmitz, & Teige-Mocigemba, 2007). Recent research has shown that measures that pretend to assess bottom-up processes can be influenced by higher cognitions (e.g.,Alexopoulos, Lemonnier, & Fiedler, 2017). However, Fiedler and Hütter (2014) pointed out that the limits of automaticity also have their limits in the sense that under extreme conditions paradigms could be constructed in such a way that nearly all controlled influence is eliminated.

Knowing about the difficulties of interpreting an indirect measure, we nonetheless think that they could provide insights into the evaluative processing of ambiguity beyond what direct measures can provide. Based on the generality of automatic activation of evaluations, we hypothesize that a person without deliberately processing the two meanings of the ambiguous information, has a fast, unintentional activation of a positive or negative evaluation which some research may define as an affective reaction or feeling.

We used the EP paradigm to assess the automatic activation of evaluations of ambiguity/clearness and the IAT to assess indirectly the attitude towards ambiguity/clearness. In the following sections we will describe the procedure, the logic behind the two measures and their strengths and weaknesses in order to make clear why we chose these two measures the way we used them (IAT as moderator of the EP paradigm). Finally, we will point out why it is necessary at all to include a moderator of the effect in the EP paradigm in order to answer the main question: Are there interindividual differences in the automatic activation of evaluation of ambiguity/clearness?

Affective/Evaluative Priming. Affective/evaluative priming is the phenomenon that affectively/evaluatively positive or negative primes facilitate reactions to targets that share the same valence. The basic logic is that the prime is evaluated automatically and that the evaluative congruency with the target leads to faster or slower responses (Fazio et al., 1986). By virtue of the sensitivity to spontaneous automatic evaluations, the evaluative priming paradigm is seen to predict spontaneous behavior (De Houwer, 2003b).

The reaction to the target could be the mere pronunciation of the target (e.g., Hermans et al., 1994, Exp. 2) or, most common in literature, the evaluative classification of the target on a good/positive – bad/negative dimension like in the seminal paper provided by Fazio et al. (1986). This task is usually referred to as the evaluative priming (EP) paradigm. There are two other versions of sequential priming paradigms in order to investigate the affective priming effect. In the lexical decision task (see, Neely, 1977, 1991) participants have to decide which target is a real word and which target is not a word. In the semantic categorization task targets have to be categorized in terms of semantic dimensions (e.g., person vs. animal; De Houwer, Hermans, Rothermund, & Wentura, 2002; Experiment 1). Meta-analytic results showed that the affective priming is stronger in the evaluative priming paradigm than in the pronunciation task and that affective priming does not occur in the other two mentioned priming paradigms (Herring et al., 2013).

In order to assess the attitude towards the primes without voluntary control or strategic influences like self-presentation, the interval between the onset of the prime and the onset of the target called stimulus-onset asynchrony (SOA) is shorter than or equal to 300 ms (Fiedler & Hütter, 2014). Although given the evidence that the evaluative priming effect is more stable or more pronounced at SOAs below 300ms (Hermans, De Houwer, & Eelen, 2001; Klauer & Musch, 2003; Klauer, Teige-Mocigemba, & Spruyt, 2009) we chose an SOA of 300ms at the upper boundary of automaticity. The reason for this was that processing an ambiguous word is more complex than an unambiguous word. The activation of representation of both meanings of an ambiguous word takes time (Simpson & Burgess, 1985) see chapter 2.2. Therefore, an SOA of 300ms with ambiguous primes probably acts like a much shorter conventional SOA because the prime is not the presented word itself, but the mental state constituted by the representations of two distinct meanings of the word. The constitution of this mental state takes time.

In the literature, there are plenty of theoretical accounts to explain the affective priming effect in the evaluative priming paradigm (see, Alexopoulos et al., 2017; Ihmels, Freytag, Fiedler, & Alexopoulos, 2016; Klauer et al., 2009; Voss, Rothermund, Gast, & Wentura, 2013). We will explain the two main classic theoretical accounts shortly and consider the implications of these accounts in terms of applying the EP paradigm in order to study the automatic evaluation of lexically ambiguous words. The first account stems from Fazio et al. (1986) who argued that the mere presentation of an attitude object (in this case the prime) automatically activates evaluations of that object which spread in the memory network. These evaluations could be congruent or incongruent with the evaluation of the target and therefore could facilitate or hinder the processing of the target. Another prominent explanation of the EP effect is response competition (Klauer, 1997; Klauer, Musch, & Eder, 2005; Klinger, Burton, & Pitts, 2000). Prime and Target are not only congruent or incongruent in terms of their valence but also in terms of their response. The underlying logic of this account is that the prime automatically activates a specific response. When the target requires the same response, responses to the target are facilitated. When the two responses are different and compete against each other responses to the target are slowed down (Klauer & Musch, 2003). Herring et al. (2013) conclude that both postulated processes indeed seem to contribute to the EP effect under specific circumstances. In contrast, Voss et al. (2013, Exp. 1b and Exp. 2a) applied a diffusion model data analysis to EP paradigms and provided evidence that the EP effect is exclusively driven by processes at the response execution stage and not due to improved identification/processing of the target. It is important to note that the first initial step in both accounts is the same, namely the automatic activation of evaluation. The difference of these two accounts lies in what is proposed after this initial automatic attitude activation: encoding facilitation of the target vs. response activation (Fazio, 2001).

Although there is a debate about the underlying process, we can state that, based on the prime target interaction in the EP paradigm, one could infer about the attitudes towards the primes. For example, if a prime type leads to faster categorization of positive targets, one could assume that the person holds a positive attitude towards this prime (De Houwer, 2003b). However, for validating indirect measures there is often an a priori assumption about the attitude concepts on a group level (Teige-Mocigemba, Klauer, & Sherman, 2010). We, on the other hand, do not know the attitude towards ambiguity and clearness on a group level and the problem is that the EP paradigm is not reliable enough (e. g. Cronbach's alpha of .64 for response-window evaluative priming, W. A. Cunningham, Preacher, & Banaji, 2001; e. g., split-half correlation of .04 (n.s.), Olson & Fazio, 2003) to draw conclusions on an individual level. Instead, based on the evidence of direct measures, we could assume that there is a huge variance in liking ambiguity and clearness in the population. However, assuming that there is no general positive or negative attitude towards ambiguity or clearness on a group level, there will be no affective priming effect (no prime x target interaction). A non-significant prime × target interaction in a sample overall does not allow to draw the conclusion that there is no affective priming effect on an individual level. Assuming that half of the sample likes ambiguity and the other half dislikes ambiguity, it results in two asymmetrical prime × target interactions. Merging the two sub-samples masks the prime × target interaction. Therefore, if the goal is to find evidence if there is an automatic activation of evaluations of ambiguity/clearness, we need information about the attitude towards ambiguity/clearness for each participant. As the EP paradigm rather reflects an automatic process, it seems to be the best way to asses the attitude towards ambiguity/clearness via another indirect measure. We think the IAT is the best choice to do that. To our knowledge, no affective priming study has been published using primes that systematically differ in the dimension ambiguousness vs. unambiguousness, yet.

IAT. The implicit association test aims to assess the attitude towards two concepts in their relation to each other (such as ambiguity vs. clearness) by measuring the strength of associations with positive and negative valence of these concepts. Unlike earlier indirect attitude measures, the IAT does not rely on sequential priming resulting in congruent and incongruent trials but on congruent and incongruent tasks/blocks. The comparison of the performance (typically the mean reaction time) of the two tasks, resulting in the IAT effect, is indicative of the relation of the attitude towards the concepts. The congruency of the tasks is established via classifying the concepts of interest (the two target categories) and evaluative polarized attributes (the attribute categories) with the same two response keys within one task. Considering the case of holding a negative attitude towards ambiguity and a positive attitude towards clearness then, the response mapping of the concepts with the attributes is evaluative congruent for a person when ambiguity + unpleasant share one response key and clearness + pleasant share the other response key or evaluative incongruent if clearness + unpleasant share one response key and ambiguity + pleasant share the other key. The basic logic behind the task is that it is easier (more difficult) to answer with one response key if the two concepts assigned to that key are strongly (weakly) associated (Nosek et al., 2007).

The relative nature of the IAT does not allow to draw any conclusions about the absolute attitude towards the target concepts. Even if there was a difference in performance in the congruent and incongruent task resulting in an IAT score deviating from zero, both concepts could still be liked or disliked. A non-zero IAT score simply indicates that one target concept is more liked or less disliked than the other target concept. Based on our results on direct measures we know that the attitudes towards clearness and ambiguity are highly negatively correlated. This indicates that when a person likes clearness, they usually dislike ambiguity. Furthermore, two attitude scores are on average both close to the neutral scale point. Therefore, in the case of applying the IAT to ambiguity and clearness, we can assume that a non-zero IAT score reflects a positive attitude towards one target category and a negative attitude towards the other target category. Although the IAT is the most prominent indirect measure of attitudes, there are a lot of open questions.

It is not entirely clear to which degree the attitude towards the target concepts is measured (De Houwer, 2001, 2008) and to which degree the attitude towards the individual target stimuli representing the target concept is measured (Govan & Williams, 2004). The results from J. P. Mitchell et al. (2003) provide evidence that the IAT is affected by both variations of individual stimuli (when holding the labels constant) and the labels of target concepts used (when holding the individual stimuli constant). Nosek et al. (2007) agree with that and sum up that the labeling of the categories determines the interpretation of the stimuli but that the individual stimuli can affect the interpretation of the category labels as well. So, both label and individual stimuli are important and may change an IAT effect but the representations of the category level seem to contribute more to the IAT effect than the representations on the level of the individual stimuli do.

There is huge evidence that the IAT like other indirect measures of attitude does not purely asses stable stored evaluations of the target concepts in mind but can be affected by the accessibility of (new, consciously) acquired information or more generally by “activation patterns of particular aspects of a multifaceted mental representation” (Cone, Mann, & Ferguson, 2017, p. 140). For instance, results from Briñol, Petty, and McCaslin (2008, Exp. 1) show that the IAT can be affected by deliberate thinking about strong arguments concerning the target concept. The same malleability of IAT effects was shown by simply pairing target concepts with positive or negative valence prior to the IAT assessment which is interpreted to mean that the IAT score can reflect the recently learned association present in the environment irrespectively of personal evaluations (Karpinski & Hilton, 2001, Exp. 3). In a similar vein, Dasgupta and Greenwald (2001) show that the presentation of positive black and negative white exemplars reduces the usual IAT effect immediately and 24 hours prior to an IAT assessment. Even for variants of personality IATs it was shown that they measure trait and state effects (Schmukle & Egloff, 2005). Another more general view is that currently available “extra personal associations” (Olson & Fazio, 2004) which do not relate to ones one’s evaluations could influence the IAT (Han, Olson, & Fazio, 2006). These studies all provide evidence that the IAT does not necessarily reflect personal evaluations determined by the own attitude a person holds. However, Nosek and Hansen (2008) argue that the distinction between one's own and extrapersonal associations cannot be adopted to "implicit" attitudes as this requires a validation process which is not part of "implicit" cognition (see, Gawronski & Bodenhausen, 2006). They argue that all learned associations (based on direct or indirect experience) belong to us and have the potential to be activated and hence influence cognitive processing. However, they provide evidence that there is almost no independent relationship between the IAT and cultural knowledge about the attitude concept when controlling for explicit attitudes. In our case it can be assumed that the knowledge of the predominant associations between ambiguity and an evaluation in society is less salient than, for example, the evaluation of ethnic groups, making it even more plausible that an IAT with lexically (un)ambiguous words as exemplars of the target concepts tends to measure "one's own" attitude towards the target concepts. Moreover, Hahn, Judd, Hirsh, and Blair (2014) provided evidence that people have a conscious insight into their own "implicit" attitudes that goes beyond what they expect others to have (Study 3). On the one hand, this can be interpreted as the IAT reflecting own attitudes more strongly than social norms. On the other hand, their results show that the IAT does not necessarily measure unconscious inaccessible attitudes, at least for the attitude objects they chose. If the unawareness is seen as a key feature of measures of "implicit" attitude, then the results of Hahn et al. (2014) might even question the property of the IAT as an indirect measure of "implicit" attitudes.

After reviewing several process models of the IAT, Teige-Mocigemba et al. (2010) sum up that they only agree that the IAT is also influenced by factors that do not relate to the construct of interest, besides what it aims to measure. We will refer to two influential process models of the IAT.

The multinomial Quadruple Process model (Quad model; Conrey et al., 2005; J. W. Sherman et al., 2008) postulated that four distinct processes are involved when performing an IAT (and other indirect tasks too). The first process is what the IAT aims to measure: the automatic activation of association (AC). The second is the ability to detect (D) the correct response. The third process reflects the ability to overcome the bias (OB) derived from the first process and the last process (G =guessing) is a response bias which takes place when no other information (no associations and no detection of correct response) for correct response is available. Interestingly in contrast to the AC the G parameter, the D and OB parameter seem to be relatively general non-attitudinal processes meaning that they did not differ across attitude objects (Calanchini, Sherman, Klauer, & Lai, 2014).

The second influential process model of the IAT stems from Meissner and Rothermund (2013) who applied a multinomial processing tree model to the IAT in order to disentangle and estimate the probability of three different processes postulated to take place in this indirect measure. The first process which potentially takes place (but has nothing to do with associations) while doing an IAT is recoding, which means that the two binary classifications are simplified to one binary classification due to classifying every stimulus based on a single feature like salience (Rothermund & Wentura, 2004) or other features like valence (Mierke & Klauer, 2003). This recoding process can only be adopted in the compatible task. Another originally postulated process is of associative nature. That means the activation of associations between the target concepts with positive or negative valence. The last process is the identification of the correct label which should always produce the correct response. Meissner and Rothermund (2013) present over 7 experiments, covering different attitude objects, evidence for the three different postulated processes taking place in the IAT. Importantly, they could show that the parameter estimating the associative process significantly predicted behavior (Exp. 7). To sum up, although recoding processes are involved in the IAT, they conclude that the associative process plays an important role. In order to avoid recoding, a recoding free IAT in which the compatibility switch randomly between trials (and not between blocks) was presented by Rothermund, Teige-Mocigemba, Gast, and Wentura (2009).

An account covering several potential processes influencing an IAT effect is the similarity account (De Houwer, Geldof, & De Bruycker, 2005). This account aforesaid that the IAT measures the similarity between targets and attributes and that the similarity highly depends on available information. It could be the strength of associations in semantic memory (which the IAT aims to measure), perceptual features, salience asymmetries or any other information which can be used to construe a similarity between targets and attributes. Further empirical support of this account was provided by Bading, Stahl, and Rothermund (2020) whose results (IAT sensitivity to relational information acquired in a EP paradigm prior to the IAT) further challenge the IAT as a measure of strength of associations.

However, despite all the remaining uncertainties about the IAT, it also has its unique strengths. It was the first indirect measure to have an acceptable reliability in terms of internal consistency (Nosek et al., 2007). For instance, De Houwer and De Bruycker (2007) showed a split-half reliability of IAT scores between .83 and .95 over three experiments using different target concepts. The meta-analysis of Hofmann et al. (2005) reported an average reliability of the IAT of .79. This value is based on 50 studies reporting internal consistency or split-half reliability. The main result of the meta-analysis is a measurement error corrected mean population correlation of .24 (mean uncorrected $r = .19$, $N = 12{,}289$) between the IAT and self-report measures. They identified several factors influencing the relation between the two types of measures (see also, Nosek, 2005). This low shared variance leaves room for incremental validity of the IAT. Indeed, there is much evidence for this type of validity. The IAT predicts criterion variance not predicted by self-report measures (Greenwald, Poehlman, Uhlmann, & Banaji, 2009).

Because of these properties of the IAT and the fact that ambiguity and clearness represent a natural opposite we decided to choose the IAT to measure the attitude towards these concepts. To our knowledge, no IAT study has been published using concepts that systematically differ in the dimension ambiguousness vs. unambiguousness.

Structural differences of IAT and EP paradigm. The IAT, as the name suggests, aims to assess the strength of associations between target concepts and positively or negatively evaluated attributes rather than the degree of automatic activation of evaluations of individual stimuli. The EP paradigm however, comes directly from the research field of automatic affective/evaluative reactions and intuitively seems more suitable to asses these types of reactions. Indeed, this naive superficial difference of the feature has empirical support. The obvious main structural difference is that, in contrast to the sequential nature of the EP paradigm, the constructs of interest (e.g. in our case ambiguity and clearness) need to be classified explicitly in the IAT (Nosek et al., 2007). De Houwer (2003b, 2008) applied the taxonomy of compatibility tasks (see Kornblum & Lee, 1995) to the IAT and the EP in order to reveal structural differences. A task can be compatible due to the mapping of relevant (relevant S-R) and irrelevant (irrelevant S-R) stimulus features with the response or due to the mapping of stimulus-stimulus features (S-S). He presented evidence that the IAT is driven by the relevant S-R compatibility. Relevant is the feature of the stimuli that determines the correct response. In our case, the feature is the category ambiguity or clearness. Due to the assignments of the responses which could be congruent (e.g. when ambiguity and negative share a response key) or incongruent (e.g. when ambiguity and positive share a response key), the responses (left or right) obtain a clear positive or negative or mixed valence. In this example, a negative attitude towards ambiguity is assumed. De Houwer (2003b, 2008) pointed out that the relevant S-R compatibility which is established by the affiliation of the stimulus and the acquired response valence is confounded with the irrelevant S-R compatibility assuming that (as it is usual in the IATs) all exemplars of a category have a positive or negative valence. Removing this perfect confounding factor by having exemplars of opposed valence within one category, De Houwer (2001) provided strong evidence that the IAT effect is mainly due to the relevant S-R compatibility because the manipulation of the (irrelevant) valence of the exemplars did not influence the IAT effect. However, when the valence of all stimuli of a category changes (e.g. only positive insects and negative flowers), the usual IAT effect is reversed indicating that the set of stimuli influences how the superordinate target categories are construed (Govan & Williams, 2004). The EP effect, instead, is driven by irrelevant S-R compatibility and not by relevant S-R compatibility (De Houwer, 2001, 2008). The irrelevant stimulus feature in the EP paradigm is the valence of the prime which does not need to be processed in order to give a correct response. In the EP paradigm there is no manipulation of the relevant S-R compatibility because the target is indicative of the correct response and the assignments of the correct responses do not change over trials.

De Houwer (2001) concluded that the IAT therefore should reflect more the associations between the relevant target dimension (e.g. ambiguity and clearness) with the valence dimension. The EP paradigm on the other hand reflects more the global attitude towards the individual prime because here, the irrelevant valence of the prime can be congruent or incongruent with the response determined by the target. Or, in a similar vein, the IAT effect is rather driven by properties of categories and the EP paradigm is rather driven by properties of the exemplars (De Houwer, 2008). These assigned properties of the two tasks seem to best reflect our purpose of examining whether (un)ambiguous stimuli automatically trigger an evaluation (investigated via an EP paradigm) in line with the attitude towards the phenomenon itself, assessed via the IAT.

Relation of EP and IAT. Olson and Fazio (2003) conducted an IAT and an EP paradigm within one study and showed that the IAT effect and the EP effect only correlated with each other when further instruction was given in the EP paradigm to strengthen the attention on the category labels of the primes by categorizing them and counting them. They concluded that the IAT reflected the associations of category labels and the EP paradigm (without the further instruction) was more driven by the evaluations of the stimuli representing the category resulting in a low IAT-EP correlation. Even though the two measures might measure more strongly a similar construct with this procedural variation of the EP paradigm, we did not adopt this for several reasons. Firstly, more recent research showed that drawing attention to the primes in the EP paradigm could lead to a reduced EP effect which could be explained by the segmentation hypothesis (Fiedler, Bluemke, & Unkelbach, 2011). Secondly, it is not possible to consciously construe the category labels (ambiguous or clear) out of the presented words in such a short time. This might work for images of black and white faces but not for ambiguous and non-ambiguous words (deciding if a word has only one meaning can only be done with a high level of uncertainty). Thirdly, we are interested in the unintentional activation of evaluations of ambiguity and clearness on an exemplar level and not on a category level. Therefore, making the EP paradigm more IAT-like, although this variation might reduce the impact of construct-irrelevant features of the prime (Olson & Fazio, 2003), is not an option in our case. But how are these two measures of attitudes related in general? The existing evidence is mixed. While for some domains studies report that the measures are uncorrelated (for self-esteem, Bosson, Swann, & Pennebaker, 2000; for condom use, Marsh, Johnson, & Scott-Sheldon, 2001; for smoking, S. J. Sherman, Rose, Koch, Presson, & Chassin, 2003), for other domains studies do provide evidence for a substantial correlation (for ethnicity r = .55 between latent variables of the measures, W. A. Cunningham et al., 2001; for female authority r = .38, Rudman & Kilianski, 2000). When and why IAT and EP paradigm produce similar results is still unclear. The mentioned different levels of representation (category vs. exemplars) may sometimes lead to different outcomes (Olson & Fazio, 2003). Furthermore, low reliability of indirect measures hides the “true” correlation which makes it difficult to find a significant substantial correlation (Nosek et al., 2007).

As mentioned before, the IAT effect is driven by evaluative associations of the target categories (which it aims to measure) and by recoding processes (Meissner & Rothermund, 2013). For the EP paradigm, no such recoding processes are postulated because there is only one relevant dimension for categorization.

Therefore, based on the reviewed differences in structure and different process models of the two paradigms, we would argue that the EP paradigm assesses more purely the automatic (unintentional and fast) activation of evaluations of stimuli in general and of exemplars representing a category in particular, compared to the IAT. As we are interested in whether there is an automatic activation of evaluations of ambiguity represented by lexically (un)ambiguous words, the EP paradigm seems to be the best indirect measure to answer this question.

As mentioned, before, we do not know whether ambiguity is generally liked or disliked more than clearness or vice versa or whether there is no general tendency in the population. Therefore, a (non-significant) interaction of prime and target in the EP paradigm is not indicative of whether an automatic activation of evaluation took place individually. However, if the prime × target interaction is moderated by the attitude towards the concepts in relation, assessed via the IAT, we have evidence that an automatic activation of evaluation took place.

3 Empirical Evidence

3.1 Database: Norms for ambiguous and unambiguous German words

To our knowledge, there is no list of German words with ratings for important (affective) dimensions especially designed for matching ambiguous and unambiguous words. Of course there are several databases with affective ratings for German words (e.g., Briesemeister, Kuchinke, & Jacobs, 2011; Kanske & Kotz, 2010; Lahl, Göritz, Pietrowsky, & Rosenberg, 2009; Schmidtke, Schröder, Jacobs, & Conrad, 2014; Võ et al., 2009). However, none of them included ratings for the meanings of ambiguous words or frequencies of meanings of ambiguous words derived from first associations. In order to compare the affective/evaluative effect of lexical ambiguous words with unambiguous words, they have to be matched for several (especially affective) properties. Otherwise, potential affective/evaluative differences found could not be attributed to the ambiguousness or unambiguousness of the words but could be a result of confounding variables. For that purpose, we developed a database with norms for ambiguous and unambiguous German words (NAUG). Therefore, data were collected for affective and other important dimensions for the ambiguous words and unambiguous words. Subsets of stimuli from this database were then used for the following studies dealing with unintentional activation of evaluation of ambiguity.

Ratings for valence, arousal and abstractness were collected for both 100 ambiguous and 100 unambiguous words and for the assumed two most frequent meanings. Additionally, word frequencies for the 100 ambiguous and unambiguous words are included in the database. Furthermore, the first association for each ambiguous word was collected in order to calculate frequencies of meanings. As collecting enough ratings from 200 words and 200 meanings for several variables is costly, we chose only three psychological variables that are assumed to be most important for our purpose out of a universe of potentially influential dimensions, besides the psycholinguistically important variable word frequency. Each participant gave ratings only for one dimension for 100 words or meanings or wrote down and classified the associations for 100 ambiguous words. Therefore, the study consisted of 13 between-subject conditions, with random subject allocation. It was aimed to collect, on average, about 30 ratings per word.

Valence and arousal. Early factor analytic studies have already provided evidence for these two affective dimensions, although originally labeled differently as ‘evaluation’ and ‘activity’, using verbal (Osgood, Suci, & Tannenbaum, 1957) and nonverbal stimuli (for an early review see, Mehrabian, 1970). These two dimensions (besides the third factor analytic derived dimension called ‘potency’) are prevalent in affective norming studies with pictures (P. Lang, Bradley, & Cuthbert, 2008), words (M. M. Bradley & P. J. Lang, 1999) or sounds (M. Bradley & P. J. Lang, 1999) as stimuli. Additionally, more recent research on the nature of affect assumes that affect is best represented by these two distinct bipolar dimensions instead of, for example, models of discrete emotional states (Posner, Russell, & Peterson, 2005; Russell, 1980, 2009). Although there is theoretical (Cacioppo & Berntson, 1994) and empirical (Berger et al., 2019) support for considering the activation of positive and negative evaluation separately, we used the common univariate bipolar conceptualization of valence dimension for reasons of comparability with other norming studies and simplicity.

Abstractness. As our conceptualization of the dimension concreteness-abstractness is strongly linked to imageability (see Appendix A1), we mentioned some important aspects of imageability. Words with a rich semantic representation are recognized faster than words with low semantic representation, and imageability is discussed as an indicator of semantic richness (Rastle, 2007). For example, there is evidence that this semantic variable influences the naming performance for low frequency words (Strain, Patterson, & Seidenberg, 1995) and the performance in the lexical decision task (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004). Imageability could be an important variable to control for in order to make ambiguous and unambiguous words similar in terms of semantic richness. This is important insofar as the number of meanings is also discussed to be an indicator of semantic richness (see, Rastle, 2007).

Word frequency. It is known that word frequency is the strongest predictor of visual word recognition (Balota et al., 2004; Brysbaert et al., 2011; Graf, Nagler, & Jacobs, 2005). The more often a word occurs in language the shorter is the time in the LDT to decide if a word is a real word or a pseudo-word. Neurophysiological data (ERP from EEG) also showed early processing differences about 150 to 190ms after word onset, depending on word frequency (Hauk & Pulvermüller, 2004). Interestingly, even in the evaluative priming paradigm the factor word frequency of the target is an important variable which can moderate (for high frequency targets even reverse) the usual EP effect (Chan, Ybarra, & Schwarz, 2006). Because of the evidence that word frequency has an important impact both on basic visual word processing and evaluative congruency effects we included a frequency estimation derived from the print-based corpus of the “Wortschatz Project” of the University of Leipzig (https://wortschatz.uni-leipzig.de/de) for each word. These frequency estimations are very similar to other existing frequency estimations from other sources and are validated by LDT data (Brysbaert et al., 2011).

3.1.1 Method

Stimuli. In a first step, ambiguous words were collected from different sources (internet, textbooks and personal communication). Additionally, the meanings according to the online dictionary DUDEN were listed. Out of this pool, 100 words which primarily have only two distinct meanings were selected. Assuming a continuum from pure homonymy to pure polysemy (Klepousniotou et al., 2008), we tried to select words which can be regarded as a homonym rather than a polyseme. Moreover, based on a priori considerations about valence, the ambiguous stimuli list should include negative, neutral and positive words. As a starting point for selecting unambiguous words, the Leipzig Affective Norms for German (LANG) database from Kanske and Kotz (2010) was used. This database includes ratings for valence, arousal and concreteness for 1000 German nouns. We classified mean ratings of the three properties for all words by dividing the scale in three equal sized intervals (low, medium or high). Then we selected 100 words so that there is an approximately equal distribution of frequencies of possible combinations. This should guarantee, for example, that it is possible to subset positively and negatively valanced words with similar arousal and concreteness levels, which could be important for some experimental approaches. Unfortunately, after data collection it turned out that one of the selected 100 unambiguous words was lexically ambiguous (the German word ‘Schimmel’ is used to refer to ‘mildew’ or ‘white horse’).

Instructions. For the rating tasks we used a modified version of the instructions which are used for the LANG database (see Appendix A1). Participants were told that they would see 100 words and that the task was to rate each word spontaneously on a 9-point scale by pressing the corresponding number-key. In order to specify the meaning of an ambiguous word, the word was presented with its meaning or domain of use e. g. ‘Bank (Geldinstitut)’ [‘bank (financial institution)’]. The instructions focused on the perceived feeling while reading the word (with meaning) in order to assess the affective instead of semantic (knowledge based) valence/arousal (see for this distinction, Itkes, Kimchi, Haj-Ali, Shapiro, & Kron, 2017). Verbal anchors for valence and arousal dimension were presented only on the instruction page. The anchors were taken from the technical report of the International Affective Picture System (P. Lang et al., 2008). The abstractness instruction focused on the imageability of the object. In the instruction for the association task (see Appendix A1) participants were told to write down the first association they had in mind and, subsequently, to select a corresponding category which corresponded to their first association on the next page. Presented categories were always the assumed two most frequent meanings of the word and a “other” category. Although there are other procedures like the rating tasks (Griffin, 1999), we chose the standard word association task (see Gawlick-Grendell & Woltz, 1994) which we extended by subsequent self-categorization in order to estimate relative meaning frequencies.

Scales. Underneath the displayed word, the numbers 1 to 9 were presented horizontally, together with the corresponding manikin (valence and arousal) from the self-assessment manikin (SAM) (see Bradley & Lang, 1994) on every second number. For the valence and arousal scale no further verbal anchor was presented during the rating phase. For the abstractness scale the endpoints ‘konkret’ [‘concrete’] and ‘abstrakt’ [‘abstract’] were presented during the rating phase.

Participants. Data were collected from November 10 to December 7, 2016. Only full data sets were analyzed. We excluded 9 participants who did not declare to speak German on native speaker level. The remaining 423 (302 female) participants (students or employees of the University of Tübingen) had a mean age of 28.82 years (*SD* = 10.02 years). They participated in exchange for course credit or could take part in a lottery for one of fifteen vouchers worth 20 euros for an online retailer. Table 1 gives an overview of participants across conditions.

Table 1

Overview of distribution and properties of participants across conditions.

		Age (in years)		
Task	*N*	*M*	*SD*	Female
Association-Classification	65	31.28	11.08	49
Valence rating (ambiguous word)	29	27.41	9.82	20
Arousal rating (ambiguous word)	27	30.30	3.18	19
Abstractness (ambiguous word)	29	28.69	9.88	21
Valence rating (first meaning)	35	26.66	7.72	23
Arousal rating (first meaning)	34	29.44	10.01	27
Abstractness rating (first meaning)	21	28.62	11.07	12
Valence rating (second meaning)	27	29.67	9.36	18
Arousal rating (second meaning)	33	28.12	10.61	21
Abstractness rating (second meaning)	21	25.23	7.74	13
Valence rating (unambiguous word)	32	28.06	8.64	25
Arousal rating (unambiguous word)	39	28.00	10.48	29
Abstractness rating (unambiguous word)	31	29.90	10.59	25

Procedure. Participants were invited to take part in the study via the mailing list of the University of Tübingen. They could directly click on the link in the e-mail to open the website and start the study, which ran in full-screen mode on the PC of the participants. Following the consent form and demographic questions, they were randomly assigned to one of the 13 different conditions. The order of word presentation for the association and the rating task was randomized. At the end, they could provide information for the lottery or for course credit. This information was not saved together with the previous data.

3.1.2 Results

A summary of variables and the statistics included in the database can be found in Appendix A2. Appendix A3 shows the results of the association-classification task. The mean ratings for valence arousal and abstractness can be found in Appendix A4 for ambiguous words and their meanings and in Appendix A5 for unambiguous words. On average 29.83 ratings were collected for each word and at least 21 ratings were available for each word. The ratings showed the known quadratic relationship between valence and arousal. For X below and above neutral valence, the arousal increased for X^2 (see M. M. Bradley & P. J. Lang, 1999; Kanske & Kotz, 2010; Võ et al., 2009). The quadratic relationship was r_{quad} = .76 for unambiguous words, r_{quad} = .50 for first meanings of ambiguous words, r_{quad} = .68 for second meanings of ambiguous words and r_{quad} = .42 for ambiguous words. All quadratic correlations were significant ($p < .001$). The mean of the two meanings of an ambiguous word correlated highly with the rating of the corresponding word for valence ($r = .71$), for arousal ($r = .77$) and for abstractness ($r = .78$). Another way of calculating a compound of the two ratings of the meanings is to incorporate the frequencies of first association of the ambiguous word. This frequency weighted mean of the two meanings of an ambiguous word correlated with the corresponding rating of the word in a similar manner for valence ($r = .72$), for arousal ($r = .79$) and for abstractness ($r = .86$). All correlations were significant ($p < .001$). Except for the abstractness dimension ($r = .23$, $p = .02$), the correlations between the ratings of the two meanings of an ambiguous word were not significant. The a priori classification of first, second and other meaning was confirmed by the association task. For each word we calculated the percentage of frequency for the a priori classified first meaning, second meaning and for the other meaning category. The average percentage over the 100 ambiguous words for first, second and other meaning was 59, 34 and 7 percent, respectively.

Reliability. The correlations of the mean ratings of the 100 unambiguous words with the ratings from the LANG database (Kanske & Kotz, 2010) for valence ($r = .97$), arousal ($r = .86$) and abstractness ($r = .95$) were all significant ($p < .001$).

Intraclass correlation (ICC). We calculated the amount of agreement of the raters over the 100 words for each dimension and for ambiguous words, unambiguous words and for each meaning of the ambiguous words separately. It could be assumed that the variance between the ratings from different raters is higher for ambiguous words (resulting in lower ICCs) than for unambiguous words, as in the first case, the raters possibly only rate one of the meanings and which meaning they pick may differ between raters. In a similar vein, we can assume that the variance between the ratings from the different raters is lower (resulting in higher ICCs) if the meanings of the ambiguous word are rated instead of the word itself. Therefore we selected the ICC (2,1) according to Shrout and Fleiss (1979) as in our case each rater rated each target and the raters were randomly selected from a population and the unit of analysis was not averaged ratings. ICCs can be found in Table 2.

Table 2

Intraclass correlations.

	Unambiguous	Ambiguous	First meaning	Second meaning
Valence	.66	.23	.26	.28[b]
Arousal	.27	.10	.11	.12
Abstractness	.47	.28	.35[a]	.32

Note. All ICCs differ significantly from 0 ($p < .001$). All ICCs for unambiguous words are significantly ($p < .001$) higher than the corresponding ICC for the ambiguous words. [a] ICC ($p = .02$); [b] ICC ($p = .03$) are significantly higher than the corresponding ICC for ambiguous words.

3.1.3 Discussion

The aim of this study was to develop a database including ratings of valence, arousal and abstractness for unambiguous words and ambiguous words as well as for their meanings together with the distribution of first associations for ambiguous words. The high correlations for the three dimensions between ratings from this study with ratings from the LANG database support the reliability of the data, although instructions and location (online vs. laboratory) differed across studies.

For valence and arousal, the uncorrelatedness of the ratings of the meanings of an ambiguous word can be interpreted as the isolated rating of that meaning being unaffected by the other meaning. This does not necessarily have to be the case as the ambiguous word was presented with its meaning in brackets, which leaves the possibility that the other meaning of the ambiguous word came to mind, too. The inspection of the low correlation between meanings of the abstractness ratings revealed that for two words both meanings are indeed accidentally highly abstract (‘Himmel’ [‘sky’ or ‘heaven’] and ‘Devise’ [‘motto’ or ‘foreign exchange’]). Excluding these two made the correlation not significant as well, which supports the view of isolated ratings of abstractness being unaffected by the other meaning.

The pattern of ICCs was surprising to some extent. As expected, all ICCs for unambiguous words were significantly higher than the corresponding ICCs for ambiguous words. However only two of six ICCs of the separate meanings were indeed significantly higher than the corresponding ICC for ambiguous words. This was unexpected as the disambiguation by explicitly referring to one meaning was expected to reduce the variability between raters with respect to these items. If the meanings were more difficult to rate in terms of valence, arousal and abstractness and hence produced between rater variance or if the raters differed in the degree of successfully ignoring the other meaning and therefore tended to show more variance is an unsolvable question given our data.

An interesting aspect for future research is to look for the moderation of TA for the amount of integration of both meanings when giving a rating of an ambiguous word. S. T. Rydell (1966) provided evidence that for contradictory adjective-noun combinations (i.e. beautiful abortion) high TA participants integrated the two components more when giving a rating compared to low TA participants who gave more extreme ratings indicating that a selection of one component took place.

3.2 Development of a direct measure of attitude towards ambiguity

Given the limitations of existing direct measures of tolerance of ambiguity outlined in section 2.3.1, the objective of this investigation was to develop a direct measure of attitude towards ambiguity which is built on a narrow definition of ambiguity (see above). While creating a new item pool, we especially included several considerations from earlier work which have not been analyzed together. Our items reflect the distinction of ambiguity as the central aspect of a situation (like in abstract art) or as a side effect (see S. T. Rydell & Rosen, 1966). We included items which cover ambiguity in the domain of art. Norton (1975) was the first who included several items designed to assess the tolerance towards ambiguity in art. He did not report factor analytic results that could provide insight into whether these art-ambiguity items would establish a separate factor, nor did he report on the relation of tolerance towards ambiguity in art to tolerance towards ambiguity in general. To our knowledge, no direct scale has been created to assess attitudes towards ambiguity in art. Furthermore, we were interested whether we would find further evidence for a factor reflecting attitude towards unambiguousness or certainty or “desire for certainty” (Geller et al., 1993) in addition to (a) factor(s) reflecting attitude towards ambiguity. Therefore, we designed items that cover this aspect, too.

In order to assess the attitude towards ambiguity on a broad basis of evaluative responses, we followed the logic of decomposition of evaluative responses into cognitive, affective and behavioral classes (Rosenberg et al., 1960). While doing so, we tried to assess diverse evaluative responses without referring to the consideration of three distinct components of attitudes (see Bagozzi, 1978; Breckler, 1984). The decomposition of different types of evaluative information is still prevalent in recent definitions of attitudes like, for instance, by Maio, Haddock, and Verplanken (2018, p. 4) who defined an attitude “as an overall evaluation of an object that is based on cognitive, affective and behavioral information".

We attempted to validate our scales and answer the question if there are substantial interindividual differences in evaluations of ambiguity via the assessment of the evaluation of presented ambiguous stimuli with parallelized unambiguous stimuli (e.g. sentences). Substantial means that the evaluations are not only random variation but correspond to the assessed attitude towards ambiguity via direct measures. An overview of the steps of the scale development can be found in Table 3.

Table 3

Overview of steps of scale development.

Step	Description	Results	Data source	*N*
1	Identification of possible types of situations in which ambiguity occur via a survey (qualitative analysis)	Unresolved ambiguity can cause strong discomfort especially in job context and other social interactions	Other Study	49
2	Exploratory factor analysis on large item pool (61 items: 49 new designed items and 12 items form established scales)	3 factor structure: 1. disliking of ambiguity in general 2. disliking ambiguity in art 3. desire for clearness	Study 1	755
3	CFA of a model with 9 items per factor	9 item per factor model has a poor fit; creating an improved model with less items (good fit indices)	Study 2	279
4	(1) CFA of improved model from step 3 (2) Assessing concurrent and discriminant validity of the 3 scales via other explicit scales (3) Assessing criterion validity via mood change after conducting and clear vs unclear task	(1) Good fit of improved model (2) Evidence for both types of validity (3) Disliking of ambiguity in general and desire for clearness predicts mood change after clear vs. unclear task	3 other studies (assessed at the same time, direct measures were at the end of all 3 studies) one of these studies (N = 190) was designed to test criterion validity	591
5	CFA on an extended version (for increasing the reliability) with 7 items per factor	7 item per factor model has a poor fit; selecting a reduced version with similar good fit indices like the improved version from step 3 but with higher reliability and more diverse content	Study 3	443
6	Using the slightly adopted version of scales from step 5: (1) Validation of scale 2 (disliking of ambiguity in art) via testing the moderating role of scale 2 for the evaluation of 2 different art types (abstract vs. representational) (2) Assessing concurrent and discriminant validity	(1) Evaluation of abstract and representational paintings was moderated by scale 2 (but also by other scales) (2) Evidence for both types of validity	Other study	55

Note. Study 1, Study 2 and Study 3 are part of this thesis and procedural details and information about the sample can be found in the corresponding chapters

3.2.1 Step 1

Method. Like Reis (1996), we started the scale development by asking participants about ambiguous situations and how they react and feel while being in these situations. These descriptions were used in step 2 in order to design items. However, in this pretest we gave participants a clear definition of what we meant by ambiguity (see Appendix B1). In particular, this pretest was supposed to reveal the domains in which ambiguity is prevalent. Data were collected between November 29, 2016 to January 26, 2017. The link to the study was distributed in social media platforms. We only analyzed data sets from participants who finished the study and responded to all questions. We collected full data sets from 49 participants. They (40 females, 9 males) had a mean age of 35.19 years (*SD* = 12.97 years).

Results. Because of the small and not representative sample the responses were analyzed only qualitatively. Nevertheless, we could infer some aspects about how ambiguity is perceived in different domains. A small selection of responses can be found in Appendix B2. (1) When ambiguity arose due to multiple meanings of a word (lexical ambiguity) it seemed that these situations were perceived as less negative and less threatening than other ambiguous situations. This could be because these situations are only ambiguous for a short time as they usually can be resolved easily by context or inquiry. In retrospect, these situations are often perceived as amusing. (2) Ambiguous instructions or comments about the executed work from a supervisor seemed to be resolved infrequently and hence produced negative rumination. (3) Short text-based communication via messenger services has the potential to be ambiguous as disambiguating information via facial expressions or gestures is missing. (4) Ambiguity could arise due to the lack of clarity as to whether a statement should be meant to be ironically or not. These situations can cause strong discomfort especially if the perceiver of the statement did not dare to resolve the situation by inquiry.

3.2.2 Step 2

Method. In this step, based on the qualitative analysis of data from step 1 and in consideration of contributions of decomposition of the evaluative response (Rosenberg et al., 1960), the distinction of ambiguity as central or noncentral for a situation (S. T. Rydell & Rosen, 1966) and the factor “desire for certainty” (Geller et al., 1993) a large item pool was generated. The Item pool consisted of 49 new items and 12 items from established scales which are compatible with our definition of ambiguity. The 61 items of the item pool can be found in Appendix B3.

They fit into a grid reflecting the domains (1) in generally social related (2) job related (3) art related (4) domain unspecific and the three classes of evaluative responses (1) cognitive/general evaluative (2) affective (3) behavioral/motivational. The classification of each item into the two-dimensional space can be found in Appendix B4. A 6-point rating scale ranging from “Stimme gar nicht zu” [“strongly disagree”] (0) to “Stimme völlig zu” [“strongly agree”] (5) was used. Higher values indicated negative attitude towards ambiguity or positive attitude towards clearness. An exploratory factor analysis on a large sample (N = 755 participants from study 1) should reveal a factor structure which is interpretable. Conceptual interpretability was our main criterion for factor retention (see Worthington & Whittaker, 2006). Although the scales based on the found factor solution were at an early stage of scale development, we investigated the correlations with other related explicit scales and the D values of two IATs which were designed to measure the relative liking and arousal of ambiguity and clearness (for details see study 1).

Results. Psychometric adequacy of the correlation matrix was assessed prior to the factor analysis (see Dziuban & Shirkey, 1974). Both the Bartlett’s test of sphericity (Bartlett, 1950) provides evidence for the appropriateness for factor analysis (χ^2 = 15338.41), $p < .001$) and the Kaiser-Meyer-Olkin measure (Kaiser, 1970) with the value .92. Item statistics for the 61 items can be found in Appendix B5. A principal axis factor analysis with varimax rotation was used. We used a pairwise deletion of missing values. There were 14 factors with an eigen value larger than 1, which offers no interpretable factor solution adopting the Kaiser criterion. According to the Scree-Plot (Appendix B6) 3 factors can be considered as relevant factors. After varimax rotation, the three-factor solution offered an interpretable solution (see Appendix B7). The first factor can be considered to reflect the disliking of ambiguity in general. The second factor can be interpreted to reflect the disliking of ambiguity in art (visual art, literature and music). The third factor can be regarded to reflect the desire for clearness and unambiguousness. After rotation the 3 factors explained 28.22 % of variance together. The correlation of the three factors (only items included which load higher or equal than .40 on only one of the three rotated factors) with other explicit scales (see Table 17; correlation Table in Study 1) provides initial support for validity as they correlated as expected moderate to high with the MSTAT-2. Furthermore, they shared less variance with the conceptual different construct uncertainty tolerance than the MSTAT-2. However, the correlations of the three factors with the valence-IAT and arousal-IAT (see also Table 17) revealed only a significant correlation of factor 3 (desire for clearness or unambiguousness) with the arousal-IAT. A stronger desire for clearness and unambiguousness corresponds (as expected) to a stronger association of high arousal with ambiguity compared to the association of clearness with high arousal. However, adjusting the alpha error (regardless of method) for multiple comparisons would shift the *p* value of .03 into the range of insignificance. Therefore, only the correlations with other explicit scales can be taken as initial evidence for concurrent and discriminant validity.

3.2.3 Step 3

Method. Based on the factor structure from step 2, new items consistent with our interpretation of factors were generated and others were deleted based on low item-total correlations or multiple factor loadings (e. g item 6) so that 9 items each representing one factor (see Appendix B8) remained. The generation of new items was done in consideration of diverse evaluative responses. These 3 x 9 items built the basis for a confirmatory factor analysis (CFA) which was done on data of those 279 participants of study 2 who completed all explicit measures. The purpose of this step was to check whether the three-factor solution identified in step 2 fit the data of a new sample. Removing items with low factor loadings should further improve the scales. In order to examine the validity of the improved scales (see Appendix B9) we correlated them with other explicit measures and the IAT *D* score. Another more exploratory attempt of validation was done by correlating the scales with the mean affect ratings assessed while reading ambiguous or clear sentences (see Appendix B10) and the amount/number of selected ambiguous sentence endings (see Appendix B11). For procedural details of these tasks see the method description of part 5 of study 2. We expected that a disliking of ambiguity and desire for clearness (indicated by higher values on our scales) corresponded to negative valence/higher arousal (positive valence/lower arousal) ratings of ambiguous (unambiguous) sentences and selecting fewer ambiguous sentence endings.

Results. Data for this step comes from study 2. The original 9 item per factor solution showed an unacceptable model fit. After improving the model fit (exclusion of items with low factor loading or high multiple factor loadings and items for which the errors were highly correlated), the model showed an acceptable fit with the 3 factor solution found in step 2 (see fit indices for 3 factor model in Table 4). The remaining items of the three factors and factor loadings can be found in Appendix B9. Most importantly, the comparison of this model and a model in which the first and third factor were put together (by setting the covariance of the first factor and third factor to 1; all latent variables have a variance of 1) showed that the three factor model fit the model significantly better. For fit indices and model comparisons see Table 4. The first factor (disliking of ambiguity in general) correlated with the second (disliking of ambiguity in art) with *r* = .39 and with *r* = .65 with the third factor (liking of clearness). The art and clearness factor correlated with *r* = .38.

Table 4

Fit indices of models and model comparisons (step 3 of scale development).

Model	χ^2	df	$\Delta\chi^2$	Δdf	$p(\Delta\chi^2)$	RMSEA	CFI	NFI
Original	798.06	321	-	-	-	.07	.79	.71
2 factors	214.61	75	583.45	246	< .001	.08	.89	.84
3 factors	142.26	74	72.35	1	< .001	.06	.95	.90

Note. N = 279. χ^2 value of all models significant $p < .001$.

Validity was investigated by correlating the three scales with other direct measures and the IAT *D* scores (see Table 5).

Table 5

Summary of intercorrelations for scores of direct measures and correlations of direct measures with the D value of the valence and arousal IAT.

	1	*2*	*3*	*4*	*5*	*6*
1. *D*	.73/ .83	.00 (.95)	-.08 (.28)	.12 (.12)	.06 (.42)	.05 (.53)
2. UT	.07 (.48)	.71				
3. MSTAT 2	-.02 (.81)	.50	.86			
4. Scale 1	-.03 (79)	-.43	-.75	.83		
5. Scale 2	.07 (.42)	-.26	-.41	.31	.80	
6. Scale 3	-.08 (.36)	-.45	-.52	.52	.28	.69

Note. Horizontal correlations of the direct measures with the *D* value were calculated with the valence-IAT (*N* = 164). Vertical correlations of the direct measures with the *D* value were calculated with the arousal-IAT (*N* = 111). *N* = 279 for intercorrelations for direct measures. UT = Uncertainty Tolerance Scale. MSTAT 2 = Multiple Stimulus Types Ambiguity Tolerance Scale 2. Scale 1 – 3 contain items after model fit improvement (see Appendix B9 for items). Scales can be interpreted as: Scale 1 = Disliking of ambiguity in general. Scale 2 = Disliking of ambiguity in art. Scale 3 = Liking of clearness. All insignificant *p* values (>.05) in brackets. All other *p* values less than .001. In main diagonal odd even reliability for valence/arousal IAT *D* value and Cronbach's alpha values for direct measures.

The results of the validation by correlations of scales with valence/arousal ratings while reading ambiguous and clear sentences and a score of choosing an ambiguous sentence ending can be found in Table 6.

Table 6

Summary of correlations of the explicit measures with the affect (valence and arousal) ratings while reading ambiguous and unambiguous sentences and sentence completion task.

	UT	*MSTAT 2*	*Scale 1*	*Scale 2*	*Scale 3*
Valence ambi.	.08 (.31)	.12 (.12)	-.11 (.16)	-.15 (.06)	.14 (.07)
Valence unambi.	.03 (.66)	.11 (.15)	-.15 (.05)	-.16 (.04)	.10 (.18)
Arousal ambig.	.06 (.54)	-.01 (.88)	.05 (.59)	-.04 (.68)	-.04 (.65)
Arousal unambi.	.18 (.05)	-.05 (.60)	.08 (.37)	-.04 (.69)	-.04 (.67)
N ambiguous end	-.02 (.75)	.17 (.01)	-.20 (< .01)	-.10 (.08)	-.14 (.02)

Note. Correlations of valence rating with direct measures (N = 167). Correlations of arousal rating with direct measures (N = 112). Correlations of sentence completion task with explicit scales (N = 279). Valence (Arousal) ambiguous (unambiguous) = valence (arousal) affect rating of ambiguous (unambiguous) sentences. N ambiguous end = How often an ambiguous sentence ending was selected. UT = Uncertainty Tolerance Scale. MSTAT 2 = Multiple Stimulus Types Ambiguity Tolerance Scale 2. Scale 1 – 3 contain items after model fit improvement (see Appendix B9 for items). Scales can be interpreted as: Scale 1 = Disliking of ambiguity in general. Scale 2 = Disliking of ambiguity in art. Scale 3 = Liking of clearness. p values in brackets.

3.2.4 Step 4

Method. In this step another CFA with new data was done. We tested the 3-factor solution again but this time only with the items derived from the improved model from step 3 (see Appendix B9). Furthermore, we tested if our scales shared less variance with the big five personality factors assessed with the short version of the Big Five Inventory (BFI-K; Rammstedt & John, 2005) than the MSTAT-2 (McLain, 2009) and the Uncertainty tolerance scale (Dalbert, 1999). We used a German translation (see Appendix I) of the MSTAT-2. Data of all direct measures was assessed at the end of three different studies (not reported in this thesis) for which the data collection period lasted one month (from 28/11/2017 to 28/12/2017). Of the 614 participants who completed the studies, we excluded data of 14 participants because they participated via Smartphone, 8 participants because they declared to speak German not on native speaker level and 1 participant for whom both exclusion criteria were true. The remaining 591 participants (447 female, 141 males, 3 missing gender information) had a mean age of 28.06 years (SD = 10.70 years).

Criterion validity. According to the person × situation interaction model people have more positive (negative) and less negative (positive) affect within a situation which fits (does not fit) to their personality (Diener, Larsen, & Emmons, 1984). One of the three studies (N = 190, 145 female, 44 male, 1 missing gender information; M age 26.63, SD = 9.30) served as a validation study. This study should reveal if there was a mood shift after participating in a clear (allowing only one solution) or unclear (allowing for multiple solutions) task depending on the direct measured disliking of ambiguity (scale 1) and liking of clearness (scale 3). We used the items of the scales from the improved model of step 3. The task was to sort 9 objects presented on the screen. There were 3 circles, 3 rectangles and 3 triangles. There was a blue, yellow and red object of each form. The instruction in the unclear task was: “Create meaningful categories by arranging the objects in groups by clicking and dragging. A group is defined by at least 2 objects.” The instructions in the clear task was: “Create three categories by arranging the objects by clicking and dragging. Sort by shape/color).” In the clear task condition the category defining property “shape” or “color” was counterbalanced. Mood was assessed prior to and after the task. Therefore, we used the German version of the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) by Krohne, Egloff, Kohlmann, and Tausch (1996). Based on reported factor loadings of items (Krohne et al., 1996), we splitted the 10 positive affect and 10 negative affect item sets. We used 5 positive and 5 negative items for assessing the mood before and after the task. An interaction of the task with the measured disliking of ambiguity and liking of clearness was expected. The scales were coded so that higher values indicate a positive attitude towards clearness (scale 3) or stronger disliking (negative attitude) of ambiguity (scale 1 and 2). For positive (1 *SD* above the centered mean) values (indicating stronger disliking of ambiguity than the average disliking of ambiguity of the sample) we expected that after a clear (unclear) task the mood is shifted in a positive (negative) direction. For negative values (indicating weaker disliking of ambiguity than the average disliking of ambiguity of the sample) the reversed pattern (crossed interaction) was expected. Mood change was calculated via the difference score of composite mood post and composite mood pre-rating. Composite mood scores were calculated via the difference scores between the mean of the 5 positive and the mean of the 5 negative items.

Results. Results of the CFA can be found in Table 7. Although the estimated correlation of the first factor (disliking of ambiguity in general) with the third factor (liking of clearness) was r = .77, the model comparison showed that the three-factor model fit significantly better to the data than a two-factor model in which the first and third factor are put together. The estimated correlations of first and second factor (disliking of ambiguity in art) with r = .40 and of second with third factor with r = .44 were like the estimated the factor correlations from the CFA of step 3. This strengthens the view of partially independent attitude towards ambiguity in art compared to the general attitude towards ambiguity and the attitude towards clearness.

Table 7

Fit indices of models and model comparisons (step 4 of scale development).

Model	χ^2	df	$\Delta\chi^2$	Δdf	$p(\Delta\chi^2)$	RMSEA	CFI	NFI
2 factors	271.95	75	-	-	-	.07	.93	.91
3 factors	207.19	74	65.76	1	< .001	.06	.95	.93

Note. N = 590. χ^2 value of all models significant $p < .001$.

The factor analytic results correspond to the correlation pattern with direct measures of uncertainty tolerance (UT), tolerance of ambiguity (TA) assessed with the MSTAT-2 and personality traits (see Table 8). The second scale (disliking of ambiguity in art) is less correlated with UT and with TA than the first and third scale. However, this scale was the only one that had a large correlation with the big five trait openness.

Table 8

Intercorrelations of scales in development, UT, MSTAT-2 and big five personality traits.

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1. UT	.73									
2. Scale 1	-.43 ***	.85								
3. Scale 2	-.27 ***	.33 ***	.80							
4. Scale 3	-.44 ***	.59 ***	.31 ***	.65						
5. MSTAT-2	.54 ***	-.75 ***	-.38 ***	-.56 ***	.87					
6. Extraversion	.28 ***	-.27 ***	-.16 ***	-.18 ***	.32 ***	.84				
7. Agreeableness	.15 ***	-.07 (.09)	-.09 (.04)	-.10 (.01)	.13 **	.18 ***	.64			
8. Conscientiousn.	-.04 (.30)	.06 (.14)	.07 (.09)	.04 (.31)	.10 (.01)	.19 ***	.04 (.32)	.73		
9. Neuroticism	-.33 ***	.34 ***	.03 (.54)	.28 ***	-.44 ***	-.29 ***	-.17 ***	-.19 ***	.81	
10. Openness	.24 ***	-.21 ***	-.60 ***	-.17 ***	.25 ***	.14 ***	.07 (.10)	-.03 (.41)	.09 (.02)	.73

Note. N = 590 (one participant was removed as he finished the study but did not respond to any BFI-K questions). *** = $p < .001$; ** = $p < .01$. All other p values in brackets. Cronbach's alpha in main diagonal.

Criterion validity. We regressed in three different models the mood change score on the interaction of task instructions (clear [shape], clear[color], unclear) with the mean centered attitude scores of the three different scales from the improved model of step 3. Two different contrast variables were used to compare the 3 different task instructions. The first contrast variable was coded so that unclear task instruction was contrasted with both versions of clear task instruction (2, -1, -1). The second contrast variable was coded to contrast the (clear [shape] with the clear[color] task instruction condition (0, -0.5, 0.5).

Table 9 shows the fixed effects of the model using scale 3 (liking of clearness) as a moderator. As predicted the attitude towards clearness significantly moderated the influence of the task condition (clear vs. unclear) on the mood change score. Figure 1 visualizes this interaction.

Table 9

Fixed effects of linear model, Scale 3 (step 4 of scale development).

	b	*SE b*	*t*	*p*
Intercept	0.94	0.08	12.47	< .001
Contrast 1	-0.05	0.05	-0.96	.34
Contrast 2	0.02	0.20	0.12	.91
Scale 3 (Clearness)	-0.02	0.09	-0.19	.85
Contrast 1 × Scale 3 (Clearness)	-0.14	0.06	-2.37	.02
Contrast 2 × Scale 3 (Clearness)	0.14	0.24	0.61	.54

Note. DV: Mood change score; Scale 3 is centered by the mean (2.88), *SD* = 0.84.

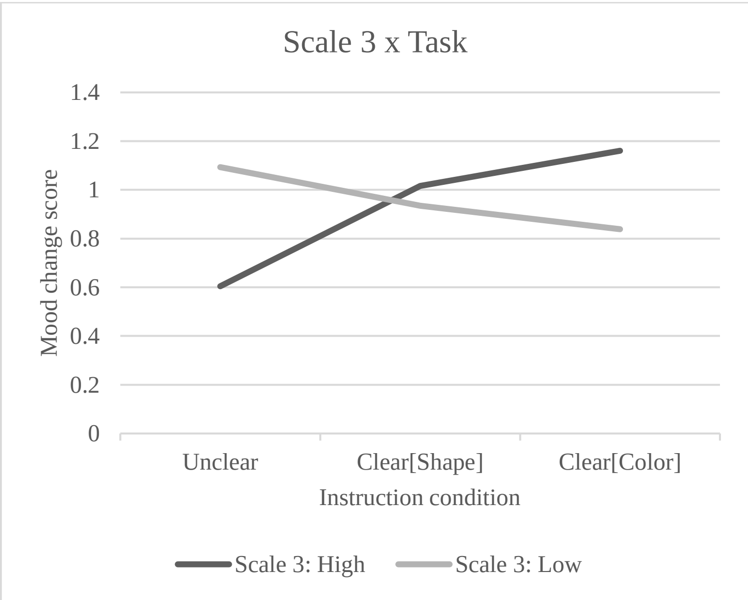


Figure 1. Predicted mood change score values for (in relation to the mean attitude towards clearness of this sample) higher and lower scores of the attitude towards clearness scale (1 *SD* above and below the mean centered scale 3) depending on the instruction condition.

Figure 1 shows that as expected the mood after conducting the task with an unclear instruction (allowing for multiple solutions) was more positive for participants which had a lower desire for clearness compared to participants with a higher desire for clearness. That was reversed for participants which received a clear instruction.

Similar results (see Table 10 and Figure 2) were found when scale 1 (disliking of ambiguity) was used as a moderator. As expected, there was no interaction and no significant main effect when scale 2 (disliking of ambiguity in art) was used as a moderator.

Table 10

Fixed effects of linear model, Scale 1 (step 4 of scale development).

	b	*SE b*	*t*	*p*
Intercept	0.95	0.07	12.82	< .01
Contrast 1	-0.05	0.05	-1.03	.30
Contrast 2	0.08	0.20	0.40	.69
Scale 1	-0.05	0.10	-0.55	.58
Contrast 1 × Scale 1	-0.12	0.06	-2.07	.04
Contrast 2 × Scale 1	0.43	0.27	1.62	.11

Note. DV: Mood change score; Scale 1 is centered by the mean (2.22), *SD* = 0.82.

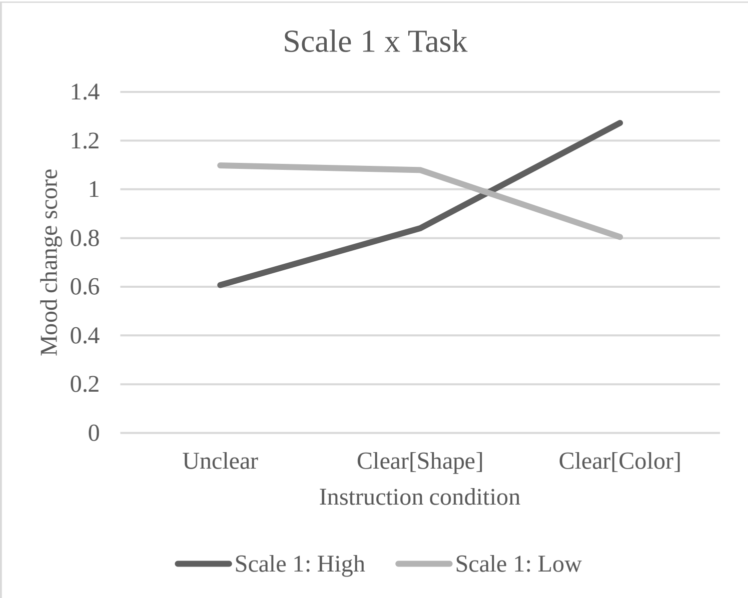


Figure 2. Predicted mood change score values for (in relation to the mean attitude towards clearness of this sample) higher and lower scores of the disliking of ambiguity scale (1 *SD* above and below the mean centered scale 1) depending on the instruction condition.

3.2.5 Step 5

Method. In this step we added new items to the improved model derived from step 3 (see Appendix B9) in line with the interpretation of the factors so that each scale now consisted of 7 items. The established and new items can be found in Appendix B12. These 3 x 7 items built the basis for a new CFA. Analysis was done with the 443 participants of Study 3 (347 females, 91 males, 5 missing gender information) with a mean age of 25.81 years (*SD* = 8.38 years).

Results. Results of the CFA of the extended solution with 3 factors having 7 items each (see Appendix B12), a reduced (based on factor loadings) model with 3 x 5 items (see bold printed items in Appendix B12) and the improved model of step 3 (see Appendix B9) can be found in Table 11. Although the solution from step 3 showed the best fit, we selected the 3 x 5 solution for further development as it extended the scales 2 and 3 compared to the improved model from step 3, thus resulting in more diverse scales in respect to content. This 3 x 5 item solution showed an acceptable fit. Cronbach's alpha for the first scale was .82, for second scale .86 and for the third scale.73. Especially scale three showed an improvement in reliability estimation compared to the version of the improved model from step 3. For the 3 x 5 item solution factor 1 correlated with factor 3 with r = .79, with factor 2 with r = .39. Factor 2 correlated with factor 3 with r = .41.

Table 11

Fit indices of models and model comparisons (step 5 of scale development).

Model	χ^2	df	$\Delta\chi^2$	Δdf	$p(\Delta\chi^2)$	RMSEA	CFI	NFI
3 x 7	518.93	186	-	-	-	.06	.90	.86
3 x 5	222.06	87	296.87	99	< .001	.06	.95	.91
Step 3	160.67	74	61.39	13	< .001	.05	.96	.93

Note. N = 443. χ^2 value of all models significant p < .001.

3.2.6 Step 6

Method. This step served as validation, especially for scale 2 (disliking of ambiguity in art). It had already been shown that there is a relationship between TA and art preference. For instance, Furnham and Avison (1997) showed that TA assessed via the scale from Mac Donald (1970) correlated with the preference for surreal paintings with few elements but neither for surreal paintings with many elements nor with representational art. They concluded that the relationship between TA and art preference is low. In order to test if scale 2 has a stronger predictive power for the art preference, we collected ratings (subjective overall positivity, subjective overall negativity, valence, arousal and abstractness of the content of the paintings) of 11 abstract and 11 representational paintings (see Appendix B13) which had been used in previous studies dealing with personality and art preference and which are neutral in terms of valence (Rawlings, Barrantes i Vidal, & Furnham, 2000). Besides these ratings, we also assessed our scales (derived from step 5), the German version of the common ambiguity tolerance questionnaire MSTAT-2 (McLain, 2009), the Uncertainty Tolerance Scale (Dalbert, 1999), a short version of the Big Five Inventory (BFI-K; Rammstedt & John, 2005), a short version (16 items) of the Need for Cognition Scale (Bless, Wänke, Bohner, & Fellhauer, 1994) and the subscale contradiction of the Dialectical Self Scale (DSS; Spencer-Rodgers et al., 2015). Data for this online study was assessed between 27/06/2018 to 17/07/2018. An *N* of 50 was aspired. We advertised the study via printed flyers which were distributed in front of the library of the University of Tübingen. All participants participated in exchange for course credit or could take part in a lottery for one of 40 vouchers over the amount of 5 Euro (accepted by many shops in Tübingen). Only participants who actively confirmed that they speak German on a native speaker level and that they did not use a smartphone for participating in the study could take part in the study. We excluded one participant who only responded to 4 of 103 items of explicit scales. The remaining 55 participants (44 females, 11 males) who fully completed the study had a mean age of 25.02 years (*SD* = 5.76 years).

Results. The 11 representational and 11 abstract paintings differed in abstractness and objective arousal but not in subjective positivity, subjective negativity and objective valence (see Table 12).

Table 12

Comparison of mean ratings of 11 abstract and 11 representational paintings.

	Abstract		Representational			
Variable	*M*	*SD*	*M*	*SD*	*t*	*p*
Subjective Positivity	53.12	14.14	56.90	13.05	-1.69	.10
Subjective Negativity	40.71	13.39	37.79	14.23	1.29	.20
Obj. Abstractness	76.97	10.38	13.07	9.10	34.95	< .001
Obj. Valence	52.54	10.98	53.20	10.33	-0.33	.75
Obj. Arousal	55.12	12.67	41.48	13.65	5.01	< .001

Note. Obj. = Objective. Paired sample t-tests were calculated; 2-tailed significance is reported. N = 55; Ratings had a potential range from 0 to 100.

We used a multilevel model with random intercepts for participants and for the 22 paintings within the rating type (positive or negative). Calculations of models were done using the *lmer* function from the *lme4*-package (Bates, Mächler, Bolker, & Walker, 2015) in RStudio (R Core Team, 2017). In the fixed effects model, we regressed the ratings on art type (abstract, representational) × rating type (subjective positive, subjective negative) × scale 2 (mean centered) value and all subordinate two-way interactions and the main effects.

The mixed-effect model was based on 2420 observations from 55 participants. Random effect estimation for variance (*SD*) for the intercepts of participants was 30.44 (5.52), for the paintings by rating type 41.04 (6.40). Fixed effects of the model can be seen in Table 13.

Table 13

Fixed effects of specified mixed-effect model (step 6 of scale development).

	b	*SE b*	*df*	*t*	*p*
Intercept	47.13	1.31	64	35.86	< .001
Rating type	-7.88	1.08	40	-7.27	< .001
Scale 2	-2.26	0.96	53	-2.36	.02
Art type	-0.21	1.08	40	-0.20	.84
Rating type × Scale 2	1.97	0.53	2319	3.73	< .001
Rating type × Art type	1.67	1.08	40	1.54	.13
Scale 2 × Art type	-0.64	0.53	2319	-1.21	.23
Rating type × Scale 2 × Art type	2.13	0.53	2319	4.02	< .001

Note. Effect coding was used for the dichotomous variables: Rating type (-1 =subjective positivity, 1 = subjective negativity), Art type (-1 = representational, 1 = abstract). Scale 2 was centered by the mean (*M* = 1.56; *SD* = 0.93).

Two different perspectives on the predicted values for the Rating type × Scale 2 × Art type interaction can be seen in Figure 3 and Figure 4.

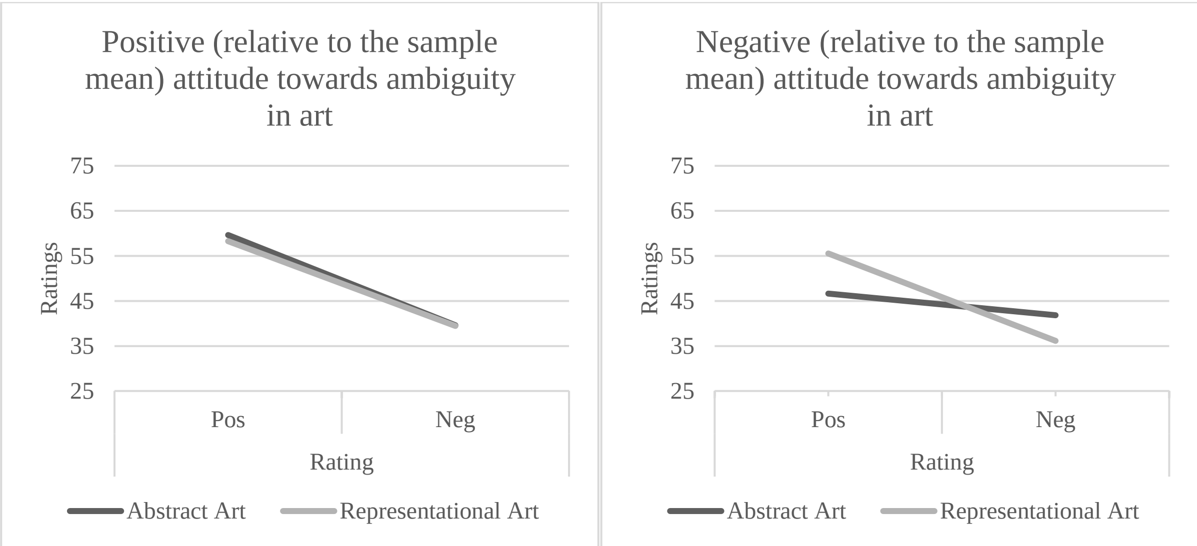


Figure 3. Visualization of the moderation of interaction of Rating type (subjective positivity and subjective negativity) and Art type by Scale 2. As Scale 2 was coded in such a way that higher values indicate disliking of ambiguity in art, higher values can be taken as negative attitude towards ambiguity in art. Left (Right) plot shows the predicted values of the Rating type × Art type interaction (using the fixed effects of the mixed-effect model from Table 13) for the mean centered scale 2 values – 1 *SD* (+ 1 *SD*).

As can be seen in Figure 3, for positive (in relation to the sample mean) attitude towards ambiguity in art the interaction of Rating type and Art type is not significant (b = -0.31, $SE\ b$ = 1.19 , t = -0.26, p = .80) but for negative (in relation to the sample mean) attitude towards ambiguity in art the interaction of Rating type × Art type was significant (b = 3.65, $SE\ b$ = 1.19, t = 3.07, p = .003). Dummy coding the variable Rating type revealed that (for participants with more negative attitude towards ambiguity in art) the influence of Art type was significant for positive ratings (*b* = -4.46, *SE b* = 1.68, *t* = -2.65, *p* = .01). This was in line with the prediction as participants with a negative attitude towards ambiguity in art rated the abstract paintings (for which the interpretation could be ambiguous) less positive than the representational paintings (for which the interpretation is probably less ambiguous). The influence of Art type for ratings of subjective negativity (for participants with more negative attitude towards ambiguity in art) was marginally significant (*b* = 2.84, *SE b* = 1.68, *t* = 1.69, *p* = .096). Descriptively these participants rated the abstract paintings more negative than the representational paintings.

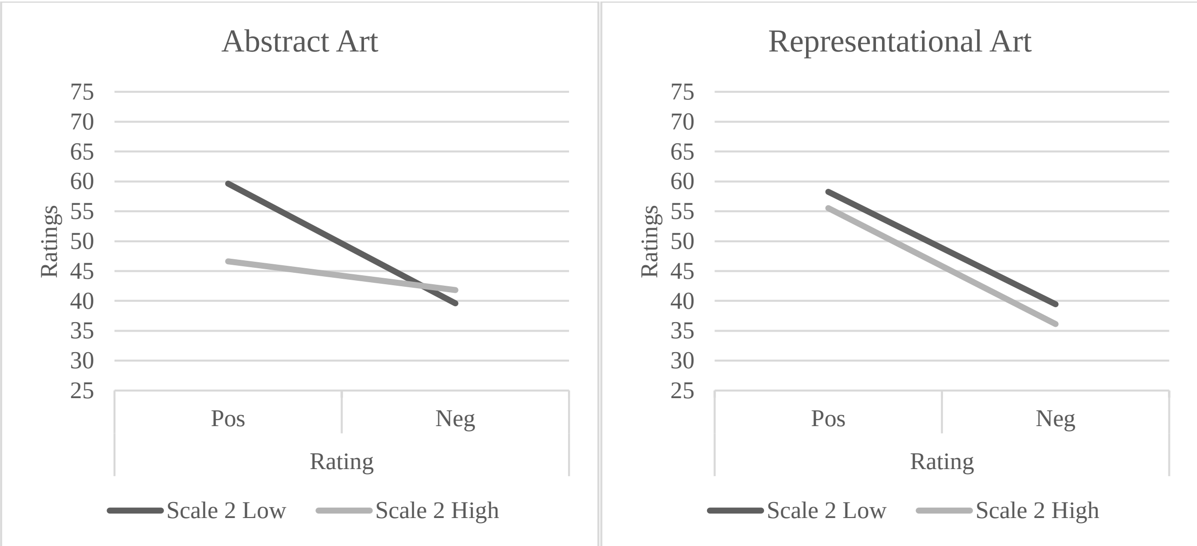


Figure 4. Visualization of the moderation of interaction of Rating type and Scale 2 by Art Type. Left (Right) plot shows the predicted values of the Rating type × Scale 2 interaction (using the fixed effects of the mixed-effect model from Table 13) for abstract (representational) Art. Scale 2 low (high) means 1 SD below (above) the centered mean. As Scale 2 was coded in such a way that higher values indicate disliking of ambiguity in art, higher (lower) values can be taken as negative (positive) attitude towards ambiguity in art in relation to the sample mean of scale 2.

Dummy coding the variable Art type revealed that the Rating type × Scale 2 interaction was significant for abstract art (*b* = 4.09, *SE b* = 0.75, *t* = 5.48, *p* < .001) but not for representational art (*b* = -0.16, *SE b* = 0.75, *t* = -0.21, *p* = .83). This means that for representational paintings the ratings of subjective positivity and subjective negativity were not differently determined by positive and negative attitude towards ambiguity in art. However, this was true for abstract paintings. For these art type, participants with low Scale 2 values (having disliking of ambiguity in art and therefore more positive attitude towards ambiguity in art in relation to the sample mean of scale 2) rated the abstract paintings more positive than the participants with high values of scale 2 (*b* = -6.99, *SE b* = 1.36, *t* = -5.28, *p* < .001). The influence of scale 2 (for abstract art) was not significant for ratings of subjective negativity (*b*

= 1.19, *SE b* = 1.33, *t* = 0.90, *p* = .37). That means that participants with a positive and negative attitude towards ambiguity in art did not differ in their ratings of subjective negativity of abstract paintings.

The same multilevel model with the same structure of fixed effect was done for scale 1, scale 3, the MSTAT-2 and the Uncertainty Tolerance Scale. Interestingly, the interaction of scale 1 (disliking of ambiguity in general) × Rating type × Art type interaction was significant (*b* = 3.26, *SE b* = 0.62, *t* = 5.23, *p* < .001), too. For a visualization of this interaction see Figure 5. Furthermore, the interaction of scale 3 (liking of clearness) × Rating type × Art type interaction was significant (*b* = 2.68, *SE b* = 0.65, *t* = 4.15, *p* < .001). This interaction is visualized in Figure 6. The interaction of MSTAT-2 × Rating type × Art type interaction was significant (*b* = -2.49, *SE b* = 0.78, *t* = -3.19, *p* = .001). The negative *b* value is due to reversed coding of this scale. The interaction of Uncertainty Tolerance Scale × Rating type × Art type interaction was not significant (*b* = -0.99, *SE b* = 0.78, *t* = -1.26, *p* = .21).

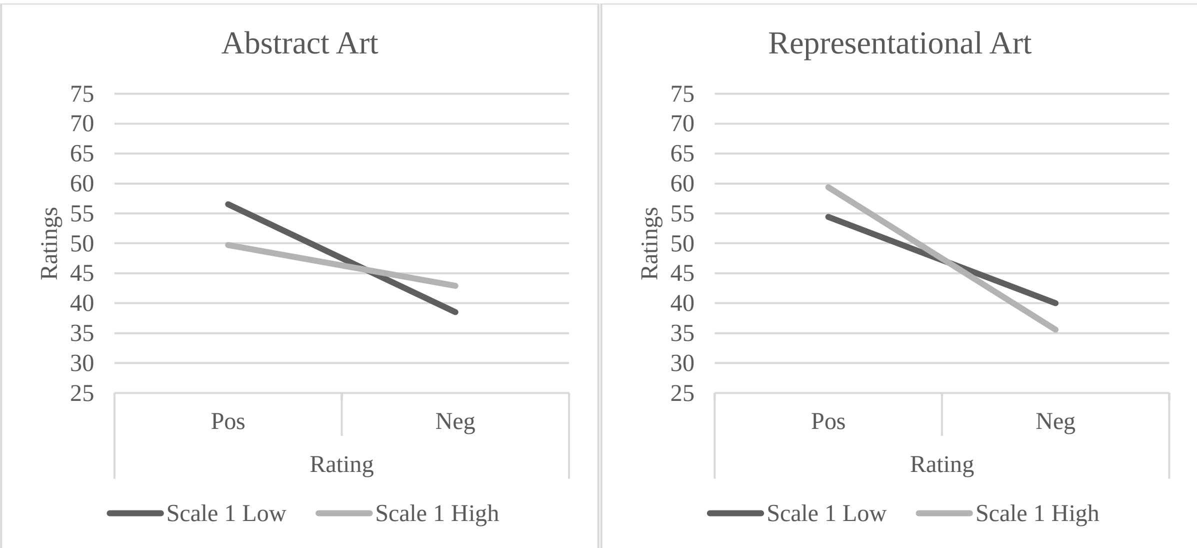


Figure 5. Visualization of the moderation of interaction of Rating type and Scale 1 by Art Type. Left (Right) plot shows the predicted values of the Rating type × Scale 1 interaction (using the fixed effects of the mixed-effect model) for abstract (representational) Art. Scale 1 low (high) means 1 *SD* (.79) below (above) the centered mean. As Scale 1 was coded in such a way that higher values indicate disliking of ambiguity in general, higher (lower) values can be taken as negative (positive) attitude towards ambiguity in general in relation to the sample mean of scale 1.

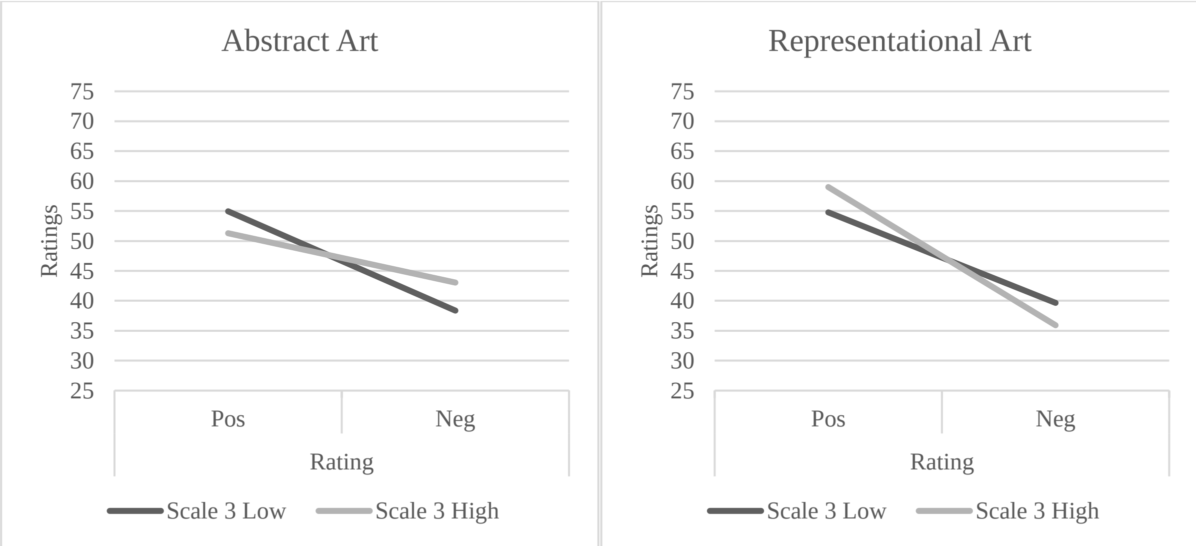


Figure 6. Visualization of the moderation of interaction of Rating type and Scale 3 by Art Type. Left (Right) plot shows the predicted values of the Rating type × Scale 3 interaction (using the fixed effects of the mixed-effect model) for abstract (representational) Art. Scale 3 low (high) means 1 *SD* (.76) below (above) the centered mean. As Scale 3 was coded in such a way that higher values indicate liking of clearness, higher (lower) values can be taken as positive (negative) attitude towards clearness in relation to the sample mean of scale 3.

Table 14

Intercorrelations of direct measures (step 6 of scale development).

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
1. Scale 1	.86											
2. Scale 2	.12 (.40)	.81										
3. Scale 3	.60 ***	.31 (.02)	.76									
4. UT	-.29 (.03)	-.12 (.39)	-.21 (.12)	.72								
5. MSTAT-2	-.78 ***	-.14 (.32)	-.58 ***	.48 ***	.89							
6. DSS	-.23 (.09)	-.31 (.02)	-.26 (.05)	.18 (.20)	.21 (.12)	.60						
7. NFC	-.36 (.01)	-.04 (.79)	-.23 (.09)	.35 (.01)	.56 ***	-.08 (.54)	.85					
8. Extraversion	-.03 (.89)	-.02 (.88)	-.17 (.22)	.20 (.15)	.24 (.08)	.18 (.19)	.29 (.03)	.80				
9. Agreeableness	.05 (.69)	-.13 (.36)	-.05 (.74)	-.09 (.51)	.01 (.92)	.23 (.10)	-.15 (.27)	.18 (.19)	.59			
10. Conscientious.	.09 (.50)	.15 (.29)	.20 (.15)	-.04 (.75)	.03 (.84)	.13 (.34)	.16 (.24)	.10 (.46)	.06 (.67)	.57		
11. Neuroticism	.29 (.03)	-.02 (.88)	.32 (.02)	-.23 (.09)	-.39 **	.19 (.17)	-.30 (.03)	-.15 (.26)	.07 (.61)	.05 (.73)	.64	
12. Openness	-.14 (.30)	-.58 ***	-.12 (.39)	.24 (.08)	.13 (.33)	.34 (.01)	.23 (.08)	.20 (.14)	-.07 (.61)	.13 (.33)	.03 (.80)	.80

Note. N = 55; *** = $p < .001$; ** = $p < .01$. All other p values in brackets. Cronbach's alpha in main diagonal.

Although the ratings of the valence and arousal of the content of the paintings should be less influenced by the individual attitude towards ambiguity, we investigated the interactions of our scales with the art type for both of these rating dimensions. Table 15 shows the fixed effects for the valence dimension and Table 16 shows the fixed effects of the arousal dimension.

Table 15

Fixed effects of the specified mixed-effect model with objective valence of the content of the paintings as dependent variable.

	b	*SE b*	*df*	*t*	*p*
Intercept	52.87	1.83	32	28.85	< .001
Art type	-0.33	1.60	20	-0.21	.84
Scale 1	-0.79	1.64	51	-0.48	.63
Scale 2	-1.04	1.17	51	-0.89	.38
Scale 3	0.47	1.78	51	0.26	.79
Art type × Scale 1	-2.79	0.84	1131	-3.32	< .001
Art type × Scale 2	-2.79	0.60	1131	-4.64	< .001
Art type × Scale 3	0.44	0.91	1131	0.49	.62

Note. Effect coding was used for the dichotomous variable Art type (-1 = representational, 1 = abstract). All scales were centered by the mean.

Table 16

Fixed effects of the specified mixed-effect model with objective arousal of the content of the paintings as dependent variable.

	b	*SE b*	*df*	*t*	*p*
Intercept	48.30	2.28	28	21.21	< .001
Art type	6.82	2.08	20	3.27	< .01
Scale 1	-0.13	1.80	51	-0.07	.94
Scale 2	-1.82	1.28	51	-1.42	.16
Scale 3	2.78	1.95	51	1.43	.16
Art type × Scale 1	-4.75	1.06	1131	-4.50	< .001
Art type × Scale 2	-1.02	0.76	1131	-1.36	.18
Art type × Scale 3	-0.65	1.14	1131	-0.57	.57

Note. Effect coding was used for the dichotomous variable Art type (-1 = representational, 1 = abstract). All scales were centered by the mean.

3.2.7 Discussion

The initial exploratory factor analysis conducted in step 2 with the entire item pool (61 items) revealed an interpretable three factor solution: (1) disliking of ambiguity in general, (2) disliking of ambiguity in art and (3) liking of clearness. Obviously, the factor structure did not reflect all domains prevalent in the item pool. The items designed to assess the attitude towards ambiguity in social situations and job-related situations did not have enough shared variance in order to establish separate factors. This could be taken as partial evidence for the domain insensitivity of ambiguity. However, the attitude towards ambiguity in art seems to be determined by a separate factor indicating that the evaluative responses to ambiguity in general and to ambiguity in art can be different. This result supports the notion of S. T. Rydell and Rosen (1966) who claimed different reactions to ambiguity depending on the centrality of ambiguity in that situation. In art, ambiguity can be seen as an inherent element, but in other stimuli ambiguity often occurs accidentally. The correlations of the first version (step 2) of the three scales with other related explicit measures and the IAT *D* scores for valence and arousal served as a first validation of the scales. As we tried to include only items in the item pool that were compatible with our narrow definition of ambiguity (see 2.1), we expected that our scales shared less variance with the uncertainty tolerance scale than the established ambiguity tolerance scale MSTAT-2 as the MSTAT-2 scale contains uncertainty as one of four types of ambiguous situations (see 2.3.1). All three scales had a moderate correlation with the uncertainty tolerance scale UT according to Cohen's conventions (Cohen, 1988). In contrast, the MSTAT-2 had a high correlation with the UT. This can be taken as first indication of evidence for discriminant validity. However, compared to the other scales of TA the MSTAT-2 is the scale which is most compatible with our definition of ambiguity. Therefore, we expected moderate to high correlation of the scales with the MSTAT-2. The correlations ranging from $r = -.48$ (scale 3) to $r = -.81$ (scale 1) can be taken as first indication of evidence for convergent validity. The correlation pattern of the explicit scales dealing with ambiguity with the valence and arousal IAT was unexpected. Based on the meta analytic results from Hofmann et al. (2005) that found a mean population correlation of .24 (uncorrected $r = .19$) between explicit measures and the IAT, we expected a low correlation between these scales with the IATs, too. This hypothesis of low correlation of our scales with the IAT was supported by evidence that the correlations with the IAT are stronger for affective explicit measures compared to cognitive explicit measures (Banse, Seise, & Zerbes, 2001) and the fact that the items of the first versions of the scales to some extent contained affective content, too. Only scale 3 showed a low significant (but only without applying an alpha adjustment for multiple comparisons) correlation with the arousal IAT in the expected direction. This result can be interpreted as evidence for no (or at most a very weak) relation between direct measures of attitude towards ambiguity with the IAT. As there is a huge variance of correlations between domains (Hofmann et al., 2005; Nosek & Smyth, 2007) it is not implausible that there actually is no relation between the IAT and direct measures for the topic of ambiguity.

Conducting a CFA in step 3 revealed a poor fit of the model with 9 items per factor (see Appendix B8). An improved model with less items was found which showed a good fit. Interestingly, although highly correlated with the first factor (disliking of ambiguity in general) the third factor (liking of clearness) seems to be an independent factor. This was supported by model comparisons which contrasted the full model and a restricted model in which the first and third factor were put together (by including a restriction that the covariance between these factors has to be 1). The three-factor model fit the data significantly better than the two-factor model. Correlations of scales of this improved model with the UT and MSTAT-2 revealed the same pattern as before. The significant low negative correlation of scale 1 and 3 with the number of choosing an ambiguous sentence ending can be taken as criterion validity. Validation by assessing the correlation with valence and arousal ratings of ambiguous and unambiguous sentences was not successful. Again, neither the MSTAT-2 nor the scales had a significant correlation with either the valence or the arousal IAT. This supports the evidence that ambiguity is a topic for which the IAT and direct measures are not correlated.

In step 4 we assessed the correlations of the scales with the big five personality traits. This revealed that scale 2 was the only scale having a large correlation with the trait openness. Moreover, scale 2 correlated less with the UT scale and with the MSTAT-2 than the first and third scale. Besides the CFA results, this correlation pattern indicates that the attitude towards ambiguity in art can be distinguished both from the attitude towards ambiguity in general, and from the attitude towards clearness. The results from the regression analysis in which scale 1 and scale 3 both significantly moderated the influence of the task condition (clear vs. unclear) on the mood change score but not the scale 2 were in line with this interpretation. This analysis served as validation for scale 1 and 3. As far as we know, no one has tested this implicit assumption (mood change after conducting an ambiguous task depending on the attitude towards ambiguity) which is prevalent in many direct measures, before. Instead, there is only evidence that high (vs. low) ambiguity tolerant (assessed with the Budner's (1962) scale) participants rate an ambiguous task as easier and more enjoyable (Ebeling & Spear, 1980).

In step 5, small changes of the length of the scales were made. Scale 1 was reduced by one item to five items and scale 2 and 3 were extended by one item to five items each. The extensions improved the reliability of the scales 2 and 3. The model fit tested via CFA was still good.

In step 6, the aim was to validate scale 2 with ratings of abstract and representational paintings. We expected that especially scale 2 would be predictive for art type preference. Results showed that subjective positive/negative overall rating of abstract and representational paintings was highly significantly moderated by scale 2 values. Unfortunately, this was true for the other two scales and the MSTAT-2 too. This, unlike earlier work (Furnham & Avison, 1997), implies that even interindividual variance in the general attitude towards ambiguity is strongly linked to art type preference. Correlation patterns of the direct measures further provided evidence for discriminant validity: The MSTAT-2 correlated more strongly with NFC, UT and the Neuroticism scale of the BFI-K than the three newly developed scales. However, in line with theoretical considerations and empirical results (Caligiuri, Jacobs, & Farr, 2000), scale 2 again was strongly correlated with the openness scale from the BFI-K. Although not directly indicative for the validation, we investigated the predictive power of the scales for the objective valence and arousal of the content of the abstract and representational paintings. Interestingly, although these dimensions should be influenced less by the individual attitude towards ambiguity, we found an interaction of art type with scale 1 for both dimensions. This indicates that the participants had difficulties rating the objective valence and arousal of the content of the abstract and representational paintings independently of their individual attitude towards ambiguity. A possible reason for that result is that the subjective valence and arousal (determined by the individual attitude towards ambiguity) triggered by perceiving an abstract or representational painting influenced the objective rating of the content of the paintings. Interestingly, for the objective arousal dimension, only scale 1 (disliking of ambiguity in general) interacted with the art type and not scale 2 or 3. This could be explained by the dominance of the affective items of scale 1 like “Mehrdeutige Situationen machen mich ein wenig nervös” [Translation: “Ambiguous situations make me a little nervous”]. For the objective valence dimension scale 1 and 2 interacted with the art type. Scale 3 (liking of clearness) did not interact with the art type when regarding the objective arousal and objective valence ratings and when the interactions of scale 1 and 2 with art type were included. This indicates that scale 3 had no additional predictive power beyond scale 1 and 2 for predicting the objective valence and arousal of the content of abstract and representational paintings.

To sum up, we established three scales with strong evidence for factorial validity. In line with previous work (Geller et al., 1993) the results of the exploratory and two subsequent confirmatory factor analyses provide evidence for a discrete factor reflecting the attitude towards clearness. Although clearness and ambiguity are natural opposites, we found evidence that the attitudes towards these constructs are best represented by two latent variables.

Our work clarifies the question of domain specificity of tolerance of ambiguity: Although we explicitly included different domains, the factor structure we found did not reflect these domains (except for art). This can be taken as evidence for a predominant domain unspecific attitude towards ambiguity. Although art is one of eight domains of the TA scale from Norton (1975) neither he nor others investigated if the attitude towards ambiguity in art established a separate factor. As mentioned before, one reason why this attitude is distinct from the general attitude towards ambiguity could be that ambiguity in art compared to ambiguity in other domains is, to some extent, a more central element (S. T. Rydell & Rosen, 1966).

The scales have good internal consistency especially considered the low number of five items per scale. Scale 1 and 2 have similar internal consistency as the much longer MSTAT-2 containing 13 items.

Based on the reported evidence for factorial, discriminant, convergent, and criterion validity for our scales, along with items that we believe to be more content valid compared to most items of other measures of TA, the scales provide an adequate measure when the goal is to assess the attitude towards the specific phenomena ambiguity, to assess the attitude towards the related construct clearness or to assess the attitude towards ambiguity in art.

3.3 Study 1

As already pointed out in section (2.3.1) an affective reaction to ambiguity is assumed by most authors dealing with the concept of TA. The definition of TA provided by Budner (1962) who defined intolerance of TA as "'the tendency to perceive (i.e. interpret) ambiguous situations as sources of threat', tolerance of ambiguity as 'the tendency to perceive ambiguous situations as desirable'" (p. 29) had a huge impact on conceptualization of TA by other researchers. His conceptualization corresponds to the general view of an aversive and appetitive motivational system which seems to be sensitive to valence and arousal of stimuli (P. J. Lang, 1995). Affect seems to be an integral part of reactions to ambiguity and is considered to vary across individuals. There is evidence that affect can be best represented by two major distinct dimensions: valence and arousal (Posner et al., 2005; Russell, 1980, 2009). One view of the functions of these two dimensions is that the valence dimension is used for evaluation and the arousal dimension amplifies this evaluation and reactions to the object (Storbeck & Clore, 2008a).

Both whether there is interindividual variation of affect (ranging from negative to positive valence and from low to high arousal) connected to ambiguity and clearness assessed with the indirect measure IAT and whether this variation is related to the variation found in direct measures was aimed to be explored with the present research. Therefore, we investigated the strength of associations of both valence and arousal with ambiguity and clearness with two IATs with different attribute classifications. One with the common attribute classification pleasantness vs. unpleasantness (valence dimension) and another with activation and deactivation (arousal dimension) as attribute concepts. Studies with modified versions of the IAT designed to assess the association of arousal with alcohol (De Houwer, Crombez, Koster, & Beul, 2004) or food (Craeynest, Crombez, Koster, Haerens, & De Bourdeaudhuij, 2008) demonstrated the applicability of arousal (instead of valence) as attribute classification dimension and provided evidence for the validity of the arousal-IAT in general.

3.3.1 Hypothesis

We expected a low correlation between the valence-ambiguity IAT and direct measures of attitude towards ambiguity (ATA) for two reasons: (1) There is the assumption that the IAT reflects more the association on a categorial/conceptual level and seems to be less driven by the automatic activation of evaluation of the individual stimuli (see De Houwer, 2001, 2008; Nosek et al., 2007). Most items of direct measures of attitude towards ambiguity (ATA) operate on an abstract conceptual level, too. (2) The known average (low) relation between IAT and self-report measures (see Hofmann et al., 2005). As some items of direct scales refer to the level of arousal when interacting with ambiguity, we explored whether there is a low correlation between the arousal-ambiguity-IAT with direct measures of ATA, too.

Furthermore, the study explored whether there is a universal attitude towards ambiguity at a group level comparable to the well-known clear preference for flowers over insects as indicated by the flower-insect IAT (Greenwald et al., 1998). In a similar vein, with the arousal IAT we can investigate whether there is a stronger association of arousal with ambiguity compared to arousal with clearness on a group level.

3.3.2 Method

Participants. Data were collected for one month from February 9 to March 9, 2017. Participants were recruited via the mailing list of the University of Tübingen (N = 682) and via spreading the link to the study on social media platforms (N = 88). Data from both sources were analyzed together. Only data sets from participants who finished the study were analyzed. From the data of these 770 participants, we excluded 10 participants who declared to speak German not on a native speaker level and 1 participant who declared to having conducted the study on a smartphone. All 11 excluded participants come from the group of participants recruited via mailing list of the University of Tübingen. It remained 759 (545 females, 213 males, missing gender information: 1) participants with a mean age of 27.34 years (*SD* = 10.05 years). 386 participants took part in the valence IAT and 373 participated in the arousal IAT. All participants participated in exchange for course credit or could take part in a lottery for one of ten vouchers of the amount of 20 Euro for an online retailer.

Design. Participants either took part in the arousal or valence IAT. The sequence of congruence was counterbalanced. We a priori defined a pairing of pleasant/deactivated with clearness as congruent. Note that only the positions of attribute labels “angenehm” and “unangenehm” [“pleasant”, “unpleasant”] (for the arousal-IAT: “deaktiviert”, ”aktiviert” [“deactivated”, “activated”]) vary between participants in order to establish the sequence of congruency. The initial position (fist block) of target concepts “mehrdeutig” [“ambiguous”] (left position) and “eindeutig” [“clarness”] (right position) was fixed.

Materials. Target stimuli used in the IAT were taken from the NAUG database (see Appendix A3 – A5). There were several criteria for selection of candidates of ambiguous stimuli from the

NAUG database. We excluded all words for which the frequencies of the first and second meaning differ significantly from an equal distribution, all words for which the valence of first and second meaning differed strongly from each other (in order to exclude ambivalent ambiguous words) and all extreme positive and negative words. We used the program Match (van Casteren & Davis, 2007) to find the five best matching items between our selected candidates of ambiguous words and unambiguous words from the NAUG database. Matching dimensions were word frequency, number of letters, valence, arousal and abstractness of the word with equal weights. Unambiguous and ambiguous words did not differ in frequency, $t(8) = -.25$, $p = .81$, number of letters, $t(8) = -.28$, $p = .78$, valence of the word $t(8) = .81$, $p = .44$, mean valence of meanings $t(8) = .10$, $p = .92$, arousal of the word $t(8) = -.06$, $p = .96$, mean arousal of meanings $t(8) = -.55$, $p = .60$, abstractness of the word $t(8) = -.26$, $p = .80$, mean abstractness of meanings $t(8) = -.05$, $p = .96$. For an overview of comparison between ambiguous and unambiguous words including *M* and *SD* see Appendix C. The used attribute stimuli for the category pleasant were “warm” [“warm”], “liebevoll” [“loving”], “glücklich” [“happy”], “gut” [“good”] and “schön” [“beautiful”]. For the category unpleasant the attribute stimuli were “kalt” [“cold”], “gehässig” [“spiteful”], “traurig” [“sad”], “schlecht” [“bad”] and “hässlich” [“ugly”]. The used attribute stimuli for the category deactivated were “entspannt” [“relaxed”], “ruhig” [“calm”], “gelassen” [“serene”], “ausgeglichen” [“balanced”] and “gemütlich” [“comfortable”]. For the category activated the attribute stimuli were “angespannt” [“tense”], “aufgeregt” [“excited”], “erregt” [“aroused”], “gestresst” [“stressed”] and “nervös” [“nervous”].

Procedure. The study ran in full-screen mode with black background on the participants’ electronic device. They were told that they would participate in two independent studies, the first one dealing with categorization of lexical information (IAT) and the second with scale development. After they finished the consent form and demographic questions, they started the IAT followed by explicit scales. Explicit scales were the Uncertainty Tolerance Scale (Dalbert, 1999), an item pool for scale development with 49 new items and 12 items from established scales of tolerance of ambiguity (see Appendix B3), a translation (see Appendix I1) of the Multiple Stimulus Types Ambiguity Tolerance Scale–2 (MSTAT–2; McLain, 2009), a translation (see Appendix I2) of the subscale “Discomfort with Ambiguity” from the Need for Closure Scale (Webster & Kruglanski, 1994b) and the subscale “Offenheit des Werte- und Normensystems” [“values”] from the personality domain “Offenheit für Erfahrung” [“Openness to Experience”] from the NEO-PI-R (Ostendorf & Angleitner, 2004). At the end they could provide information for the lottery or for course credit. This information was not saved together with the previous data.

A 7 block IAT design was used. Block 3 and 6 had 20 trials each and block 4 and 7 had 40 trials each. For all participants *mehrdeutig* [ambiguous] was presented left and *eindeutig* [unambiguous] was presented right in the first block. The left (right) response key was E (I). The sequence of congruency (congruent -> incongruent vs. incongruent -> congruent) was therefore defined by the initial positions of the attribute concepts. We followed the recommendation of doubling the number of trials to 40 in block 5 (practice of reversed target classification) in order to reduce the influence of the order effect of the compatible and incompatible block (Nosek, Greenwald, & Banaji, 2005). In order to increase the probability that the participants use the intended feature of the stimuli, that is (un)ambiguousness or valence/arousal for responding, the stimuli forming an opposite pair of one dimension (e.g. unpleasantness – pleasantness) had a different color than stimuli belonging to the other dimension. Target concepts were presented in green and attribute concepts were presented in white. After an incorrect response a red X appeared on the screen for 300 ms. Inter-trial-interval (ITI) was 250 ms. Both the arousal-IAT and the valence-IAT were identical except that labels and instructions were adjusted corresponding to the replacement of negative and positive items by high arousal and low arousal items.

3.3.3 Results

There are several scores that reflect the underlying associations strength. We decided to use the improved algorithm (Greenwald, Nosek, & Banaji, 2003, Table 4) to calculate the *D* score. Compared to the conventual score the *D* measure has less method-specific variance (Mierke & Klauer, 2003) and is not significantly correlated with cognitive abilities (Wright & Meade, 2012). The *D* score was computed so that positive values indicated that ambiguity is associated more strongly with negative valence (high arousal) and clearness more with positive valence (low arousal) than ambiguity with positive valence (low arousal) and clearness with negative valence (high arousal). In short, a positive *D* score reflects stronger liking for clearness compared to ambiguity and higher arousal for ambiguity compared to clearness, respectively. We only computed a *D* score for 752 participants because data from 7 (4 from the valence IAT) participants who responded in more than 10 percent of trials in less than 300 ms were excluded from the analysis according to the procedure of data trimming for the calculation of the *D* score (see Greenwald et al., 2003)

D score and mean latencies. The mean *D* score was 0.11 (*SD* = 0.38) for the valence-IAT (*N* = 382) and 0.10 (*SD* = 0.38) for the arousal-IAT (*N* = 370). Both the mean valence IAT *D* score ($t(381) = 5.81$, $p < .001$) and the mean arousal IAT *D* score ($t(369) = 5.06$, $p < .001$) differed from 0. For the sequence congruent -> incongruent the mean *D* score for the valence-IAT (*N* = 206) was 0.12 (*SD* = 0.37) and for the arousal-IAT (*N* = 189) 0.09 (*SD* = 0.39). For the reversed sequence the mean *D* score for the valence-IAT (*N* = 176) was 0.10 (*SD* = 0.40) and for the arousal-IAT (*N* = 181) 0.11 (*SD* = 0.38). Neither for the valence IAT ($b = -0.01$, $SE\ b = 0.02$, $t = -0.58$, $p = .56$) nor for the arousal IAT ($b = 0.01$, $SE\ b = 0.02$, $t = 0.38$, $p = .71$) the sequence of congruency (effect coded) had an impact on the *D* score.

Figure 7 shows mean latencies (after data trimming according to the improved algorithm for the *D* score) and standard errors for congruent and incongruent blocks for both sequences of congruency and for both IAT types.

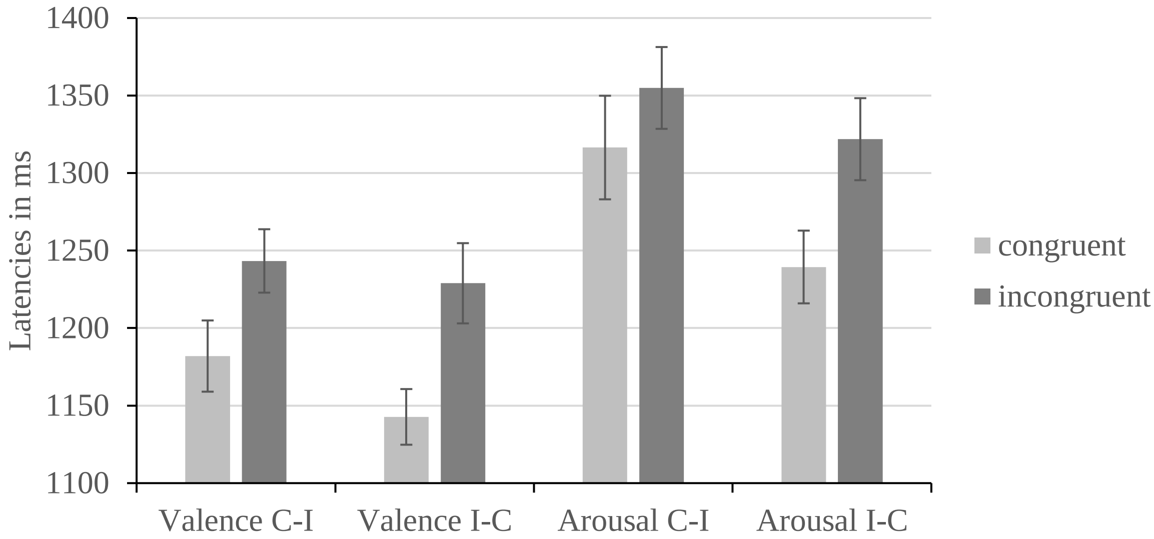


Figure 7. Mean Latencies and standard errors for congruent (ambiguous + negative/activated) and unambiguous + positive/deactivated) and incongruent (ambiguous + positive/deactivated and unambiguous + negative/activated) blocks. C-I means congruent block preceded incongruent block. I-C means the reversed sequence.

Reliability. The odd even reliability of the valence-IAT is .73 (.69 for the sequence congruent -> incongruent and .73 for the reversed sequence) and .74 for the arousal-IAT (.72 for both sequences of congruency).

Correlations with explicit measures. Data from 4 participants were excluded who missed out on so many items that at least for one scale the mean could not be computed. Data from 755 participants were correlated with the IAT *D* scores and other explicit measures. The items loading higher or equal than .40 on only one of the three rotated factors (see Appendix B7) derived from the exploratory factor analysis for the included item pool in this study (for details see section 3.2) built the scales 1 to 3. Correlations of IAT *D* values from the arousal-IAT and the valence-IAT can be found in Table 17.

Table 17

Summary of intercorrelations for scores of the explicit measures and correlations with the D value of the valence and arousal IAT (study 1).

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*
1. *D*	.73/.74	.02 (.70)	.03 (.51)	-.01 (.78)	.06 (.22)	.00 (.97)	-.09 (.07)	.02 (.69)
2. UT	-.04 (.47)	.73						
3. MSTAT 2	-.04 (.50)	.52	.88					
4. DA	.06 (.27)	-.37	-.56	.85				
5. Openness	-.02 (.71)	.23	.37	-.24	.84			
6. Scale 1	.02 (.75)	-.41	-.81	.63	-.26	.90		
7. Scale 2	.02 (.69)	-.31	-.48	.28	-.37	.38	.85	
8. Scale 3	.11 (.03)	-.32	-.50	.70	-.21	.63	.30	.77

Note. All correlations were raw correlations without applying an alpha adjustment for multiple comparisons. Horizontal correlations of the explicit measures with the *D* value were calculated with the valence-IAT (N = 380). Vertical correlations of the explicit measures with the *D* value were calculated with the arousal-IAT (N = 369). N = 755 for intercorrelations for explicit measures and reliabilities. UT = Uncertainty Tolerance Scale. DA = Discomfort with ambiguity. Scale 1 = Disliking of ambiguity in general. Scale 2 = Disliking of ambiguity in art. Scale 3 = Attitude towards clearness. *p* values can be found in brackets. Correlations without any *p* value information: p < 001. In main diagonal odd even reliability for IAT *D* value (.73 for valence-IAT), and Cronbach's alpha values for explicit measures.

3.3.4 Discussion

In this study we investigated the strength of associations of the two dimensions of valence and arousal with ambiguity and clearness using two different versions of the IAT: a valent-IAT and an arousal-IAT. Furthermore, we investigated whether these two indirect measures correlated with direct measures of attitudes towards ambiguity. Our results provided evidence that in our sample the associations of ambiguity with negative valence and high arousal and clearness with positive valence and low arousal were stronger than the associations of ambiguity with positive valence and low arousal and clearness with negative valence and high arousal. Although only a low correlation was expected, we did not find any significant correlation of direct scales with the IAT *D* scores (if an alpha adjustment for multiple comparisons was applied).

Interpretations of results. Compared to other IAT topics the *D* scores were small. For instance, a *D* score of 1.34 was reported for the young-old-IAT (Greenwald et al., 2003). Our results showed that the sequence of congruency had no impact on the *D* scores (valence and arousal) of the ambiguity-clearness IAT. This was probably due to the extra practice trials which are known to strongly reduce or eliminate the influence of the sequence of congruency (Nosek et al., 2005).

Can the *D* scores for the ambiguity-IATs be interpreted in terms of reflecting the associative strength or could other non-associative processes have contributed to the results? We know that in addition to associative processes, recoding processes also occur while conducting an IAT (Meissner & Rothermund, 2013). We can assume that the *D* score of the ambiguity-IAT is less influenced by processes of recoding based on the valence of the targets. For many IAT topics (e.g. flowers – insects) recoding the targets in the congruent block by valence seems easy to adopt since the targets strongly differ in valence. As we matched our targets by valence (and arousal, abstractness, word length, and occurrence in language), the ambiguous and unambiguous words could not be recoded easily by valence. Therefore, we assume that the targets had to be categorized in terms of ambiguity in both the congruent and the incongruent block in order to classify them correctly. If an initial processing of the ambiguousness of the stimuli is necessary, a further recoding process seems unlikely as it is not needed to respond correctly. Although valence is not the only dimension used for recoding, it is probably the most used dimension to simplify the congruent task of the IAT to a binary classification (Meissner & Rothermund, 2013).

Another potential dimension for recoding is salience (Rothermund & Wentura, 2004). It is assumed that the unpleasant/negative attribute stimuli are more salient than the pleasant/positive stimuli. If there is a salience asymmetry for the target stimuli (e.g. due to familiarity, valence), it can be used for recoding as well. Differential salience of the ambiguous and unambiguous stimuli due to familiarity or valence is unlikely because of the matching procedure. Nevertheless, there could be salience asymmetry for ambiguous and unambiguous words. Analysis of the response times for ambiguous and unambiguous words in the valence IAT revealed that the classification of ambiguous words was faster than unambiguous words (this was true for the first and third block for the valence IAT). Faster response times for ambiguous than for unambiguous targets in an ambiguity decision task probably reflect earlier termination of the search for additional meanings (when both meanings are detected) for ambiguous targets than for unambiguous targets, for which an exhaustive search is assumed (Forster & Bednall, 1976). If the analysis of response times for ambiguous and unambiguous words reflects a differential decision process with likely differential levels of certainty, this could lead to different salience of the two target groups (with ambiguity probably being more salient). This salience information could be used for recoding which could lead to faster responses in the block in which ambiguity is paired with unpleasantness resulting in a positive *D* score, which is what we found. To sum up, recoding by valence seems unlikely but recoding by salience could not be ruled out.

This potential non-associative process could lead to an increase of variance not related to the construct attitude towards ambiguity/clearness and thus could contribute to the uncorrelatedness of IAT scores with the direct measures of attitude towards ambiguity. Besides this explanation, considering potential factors determining the relationship between direct and indirect measures could help to explain why we did not find a correlation between these two types of measures for the topic of ambiguity/clearness. Nosek (2005) reported four moderator variables for this relationship: self-presentation, dimensionality, distinctiveness, and evaluative strength. Assuming low self-presentation concerns for this topic should increase the relationship. In a similar vein, the clear bipolar conceptualization of ambiguity vs. clearness should boost the relationship. We have no information to make assumptions about whether the personal attitude towards ambiguity/clearness is perceived as different to or normal among the population and hence should result in a stronger or lower relationship. However, we can assume that the attitude strength for ambiguity is rather low, contributing to a low relationship between direct and indirect measures. Furthermore, there is evidence that attitude importance is a moderator of the relationship between indirectly and directly assessed attitudes (Karpinski, Steinman, & Hilton, 2005). We can assume that the attitude importance for ambiguity/clearness is rather low too, contributing to the results we found.

A low reliability of measures attenuates the correlation and could lead to wrong conclusions about the true relationship between two variables. The reliabilities of the valence and the arousal ambiguity-IAT were both similar to the average reliability of .79 calculated based on 50 studies in the meta-analysis by Hofmann et al. (2005). Likewise, the reliabilities of the direct measures of attitude towards ambiguity could be considered as good. However, the correlations in our study were far from the reported mean uncorrected measurement correlation of $r = .19$ (Hofmann et al., 2005). Data from this meta-analysis showed that the order of indirect and direct measurements did not moderate the correlation. Therefore, too low reliability and order of presentation seems implausible as reasons for the uncorrelatedness of the indirect and direct measures. As ambiguity is not a highly socially sensitive topic, we expect that socially desirable responding is unlikely to occur for the explicit measures. Likewise, as the IAT effect is difficult to fake, although it is not completely immune to faking (Greenwald et al., 2009), we assume that this potential source of construct-irrelevant variance can be neglected when considering potential moderators of correlations between indirect and direct measures.

Whether or not the variance of the ambiguity-IAT scores indeed reflects interindividual differences in the attitude towards ambiguity cannot be answered by this study as we did not find evidence for concurrent validity. A relationship of the ambiguity-IAT together with other structurally different indirect measures like the EP paradigm (see following studies) would support the view of meaningful (construct related) variance.

3.4 Study 2

This study investigated whether lexically ambiguous and unambiguous (clear) words trigger interindividually different automatic activations of valence and arousal (as assessed with two different sequential priming paradigms) depending on the measured associations (via two different IATs) of the concepts of ambiguity and clearness with positive (low) or negative (high) valence (arousal), respectively. Activation of valence was assessed with a standard evaluative priming [EP] paradigm with ambiguous and unambiguous (clear) words as primes with neutral valence (usually, evaluative priming research employs primes with distinct positive and negative valence) and positive and negative valanced words as targets.

In the standard EP effect evaluatively positive or negative primes facilitate reactions to targets that share the same valence. The assumed underlying mechanism is that the prime is evaluated automatically and that the evaluative (in)congruency with the target leads to faster (slower) responses (Fazio et al., 1986). As already pointed out in section 2.3.2, we assumed that the evaluation of the ambiguous and unambiguous primes depends on the individual attitude towards ambiguity (ATA) and clearness. For example, for participant X (having a positive ATA) an ambiguous prime could trigger a positive evaluation, thus facilitating reactions to positive targets, while for participant Y (having a negative ATA) the same ambiguous prime could trigger a negative evaluation, thus facilitating reactions to negative targets. Therefore, we needed the IAT (which assesses the strength of associations of ambiguity and clearness with positive and negative valence) in order to use this information as potential moderator of the interaction of prime type (ambiguous or clear words) by target type (words with positive or negative valence).

Activation of arousal was assessed with an arousal priming [AP] paradigm for which targets should be classified in terms of “activating” or “not activating” (representing high and low arousal) instead of valence but in other respects identically to the standard EP paradigm. The logic of the EP paradigm could be applied to the AP paradigm. In order to pre-activate both meanings of the ambiguous words prior to the EP and AP paradigm we introduced a learning phase in which participants should learn which word was ambiguous and which word only had one meaning.

3.4.1 Hypothesis

For both priming paradigms, we expected that the prime type × target type interaction was moderated by the corresponding (valence or arousal) IAT *D* value.

Predictions for the EP paradigm: The stronger the valence-IAT indicates that the association of ambiguity with negativity (and clearness with positivity) is stronger than the association of ambiguity with positivity (and clearness with negativity), as reflected by a positive IAT *D* score, we expected a stronger prime × target interaction: if an ambiguous prime precedes a negative target, the reaction time should be faster than if an unambiguous prime precedes a negative target. If an ambiguous prime precedes a positive target, the reaction time should be slower compared to an unambiguous prime preceding a positive target. The stronger the valence-IAT instead indicates that the association of ambiguity with positive (and clearness with negative) is stronger than the association of ambiguity with negativity (and clearness with positivity), as reflected by a negative IAT *D* score, the stronger the prime × target interaction of the EP paradigm should indicate the reversed pattern.

Predictions for the AP paradigm: The stronger the arousal-IAT indicates that the association of ambiguity with high arousal (and clearness with low arousal) is stronger than the association of ambiguity with low arousal (and clearness with high arousal), as reflected by a positive IAT *D* score, we expected a stronger prime × target interaction: if an ambiguous prime precedes a high arousal target, the reaction time should be faster than if an unambiguous prime precedes a high arousal target. If an ambiguous prime precedes a low arousal target, the reaction time should be slower compared to an unambiguous prime preceding a low arousal target. The stronger the arousal-IAT instead indicates that the association of ambiguity with low arousal (and clearness with high arousal) is stronger than the association of ambiguity with high arousal (and clearness with low arousal), as reflected by a negative IAT *D* score, the stronger the prime × target interaction of the AP paradigm should indicate the reversed pattern.

A visualization of the predicted three-way interaction (here for the EP paradigm) can be found in Figure 8.

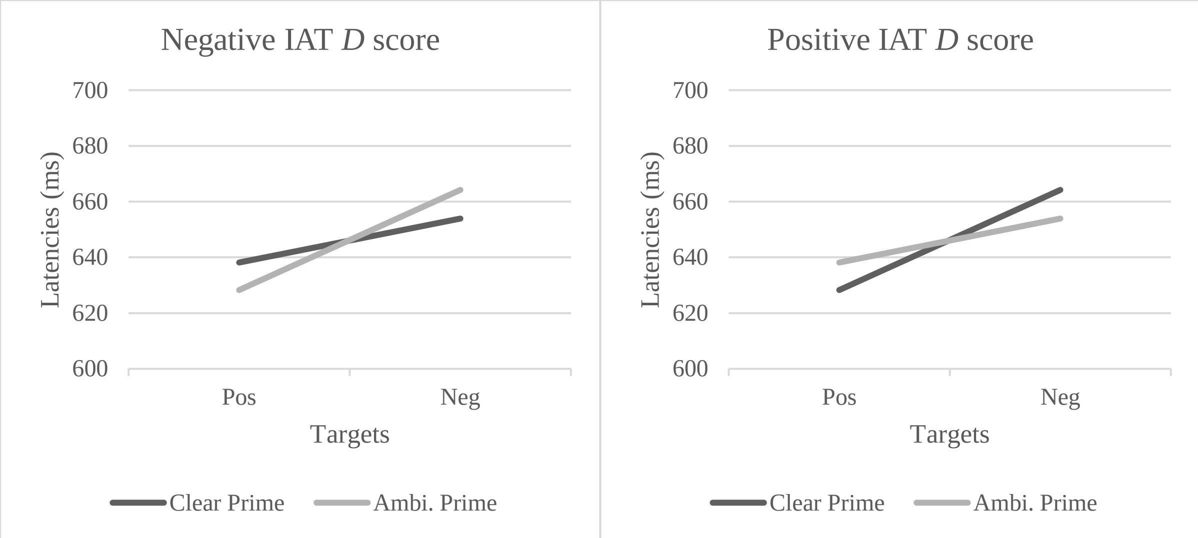


Figure 8. Predicted pattern of latencies as a function of prime type and target type and IAT *D* score. Main effect of target type was included in the prediction as it is known that the response to negative targets is slower than the response to positive targets in the EP paradigm (Dijksterhuis & Aarts, 2003; Unkelbach, Fiedler, Bayer, Stegmüller, & Danner, 2008).

3.4.2 Method

Participants. Data were collected for one month from August 10 to September 10, 2017. All participants were recruited via the mailing list of the University of Tübingen. We only analyzed data sets from participants who finished the study. Full data sets were available for 280 participants. We excluded 1 participant who declared not to speak German on native speaker level. The 279 participants (205 females, 73 males, 1 missing gender information) who finished the priming paradigm had a mean age of 27.94 years (*SD* = 10.31 years). 167 participants were assigned to the valence condition (EP paradigm, valence IAT and valence affect rating of sentences) and 112 were assigned to the arousal condition (AP paradigm, arousal-IAT and arousal rating of sentences). All participants participated in exchange for course credit or could take part in a lottery for one of ten 20 Euro vouchers of for a fair-trade shop located in Tübingen.

Design. Participants were randomly assigned to the arousal or valence condition. The same condition was used for the priming paradigm, the IAT and a task (the sentence affect rating task) designed to validate the developed direct scales (see 3.2). A repeated measure design for both priming procedures was used. For the EP paradigm the factors were 2 (prime type: ambiguous vs. unambiguous) × 2 (target type: positive vs. negative) and for the AP paradigm the factors were 2 (prime type: ambiguous vs. unambiguous) × 2 (target type: deactivated vs. activated).

Materials. All stimulus words used in the priming paradigms and the IATs were selected from the NAUG database (see Appendix A3 – A5). For the matching of word lists the program Match (van Casteren & Davis, 2007) was used. Prior to matching of ambiguous and unambiguous words we preselected ambiguous words. This selection followed several steps. First (like in study 1) we excluded all words for which the frequencies of the first and second meaning differed significantly from an equal distribution. Secondly, we excluded all words for which there were more than 10 percent of associations which did not match to the first two meanings. This was done to guarantee that all the ambiguous words used only have two frequent meanings. Then we eliminated all words for which the valence or the arousal or the abstractness ratings of the first and second meaning differed strongly from each other (absolute difference more than 2). In a last step, we excluded extremely positive and negative words. The unambiguous words and the preselected ambiguous words served as basis for the matching procedure in which the best match of 10 ambiguous and unambiguous words was identified. Like in study 1, matching dimensions (all with equal weights) were: word frequency, number of letters, valence, arousal and abstractness ratings of the word. The 10 unambiguous and 10 ambiguous words (see Appendix D2; including *M* and *SD*) did not differ in frequency, $t(18) = 0$, $p = 1$, number of letters, $t(18) = 0.82$, $p = .43$, valence of the word $t(18) = -0.67$, $p = .51$, mean valence of meanings $t(18) = 0.19$, $p = .85$, arousal of the word $t(18) = -0.02$, $p = .98$, mean arousal of meanings $t(18) = 0.39$, $p = .71$, abstractness of the word $t(18) = 0.19$, $p = .85$, mean abstractness of meanings $t(18) = 0.24$, $p = .81$.

As input for the matching procedure of positive and negative words (targets in the EP paradigm) we selected the 20 most negative and 20 most positive words from the unambiguous words from the NAUG database. Except for valence, we used the same matching dimensions as for the ambiguous and unambiguous words. The 10 positive and 10 negative words (see Appendix D3; including *M* and *SD*) did not differ in frequency, $t(18) = 0.14$, $p = .89$, number of letters, $t(18) = 0$, $p = 1$, arousal of the word $t(18) = -1.64$, $p = .12$, $p = .85$, abstractness of the word $t(18) = 0.33$, $p = .74$ but as intended in valence $t(18) = 30.27$, $p < .001$.

Input stimuli for the matching of activated and deactivated words (targets in the arousal priming paradigm) were the 31 (21) words with highest (lowest) arousal ratings from the unambiguous words from the NAUG database (unequal input stimuli size was a result of bad matchings for equal input stimuli size with 31 highest (lowest) arousal ratings). The 10 high arousal and 10 low arousal words (see Appendix D4; including *M* and *SD*) did not differ in frequency, $t(18) = 0.28$, $p = 79$, number of letters, $t(18) = 0$, $p = 1$, valence of the word $t(18) =$

-0.30, p = .51, abstractness of the word $t(18) = 0.12$, $p = .90$ but as indented in arousal $t(18) = 15.08$, $p < .001$.

Ambiguous and unambiguous words (see Appendix D5), positive and negative words (see Appendix D6) and high and low arousal words (see Appendix D7) used in the IAT are the 5 best matching pairs of the stimuli from the sequential priming paradigm. The 5 ambiguous and 5 unambiguous words did not differ (see Appendix for M and SD) in frequency, $t(8) = 0$, $p = 1$, number of letters, $t(8) = 0.96$, $p = .37$, valence of the word $t(8) = 0.29$, $p = .78$, mean valence of meanings $t(8) = 1.71$, $p = .12$, arousal of the word $t(8) = 0.41$, $p = .70$, mean arousal of meanings $t(8) = 0.70$, $p = .51$, mean abstractness of meanings $t(8) = 1.71$, $p = .13$ but in abstractness of the word $t(8) = 2.94$, $p = .02$. The 5 positive and 5 negative words did not differ in frequency, $t(8) = -0.16$, $p = .88$, number of letters, $t(8) = 0.78$, $p = .46$, arousal of the word $t(8) = -0.44$, $p = .67$, abstractness of the word $t(8) = -0.10$, $p = .92$ but as intended in valence $t(8) = 22.70$, $p < .001$. The 5 high arousal and 5 low arousal words did not differ in frequency, $t(8) = 0.12$, $p = .91$, number of letters, $t(8) = -0.41$, $p = .69$, valence of the word $t(8) = -0.20$, $p = .84$, abstractness of the word $t(8) = 0.35$, $p = .73$ but as indented in arousal $t(8) = 11.50$, $p < .001$.

Procedure. The online study used the JavaScript library jsPsych (de Leeuw, 2015). It ran in full-screen mode with black background on the PC or laptop of the participants. After confirming the consent-form they were told that the study consisted of 5 short parts which were labeled as follows: (1) Meaning of words (2) classification of lexical information (3) further development of a questionnaire (4) classification of lexical information (5) evaluation of sentences.

Following the demographic questions, participants took part in (1) the learning phase in which participants should learn which word was ambiguous and which word only had one meaning. The 20 words (10 ambiguous and 10 unambiguous) of the learning phase were the same as the primes used in the priming paradigm following afterwards. The task was to decide if the same underlined word in two sentences displayed together on the screen had one or two meanings. The same two sentences (with minimal adaption) were used for both the ambiguous and the matched unambiguous words (see Appendix D1 for the sentences used for the learning phase). This should guarantee that no differential valence or arousal transfer from the content of the sentences to the unambiguous and ambiguous words could occur and hence have an unintended effect in the priming paradigm.

Then, in part (2) the sequential priming procedure (either EP or AP) was done. There were 20 training trials (with feedback after each trial) followed by three test blocks (40 trials each) without feedback. They were told that they should answer as fast as possible. A trial consisted of a fixation cross for 500 ms followed by a blank screen for 10 ms (for technical reasons). Then the prime was displayed for 200 ms followed by 100 ms of blank screen (300 ms SOA). After that, the target was displayed until a response was given. Response keys were S (for “negative” or “not activating”) and L (for “positive” or “activating”). The ITI was 810 ms. Between blocks participants could take a break.

In part (3) the explicit measures were administered. This part consisted of the Uncertainty Tolerance Scale (Dalbert, 1999), 3 x 9 items designed to test (via confirmatory factor analysis) the three factor solution found in study 1 (see section 3.2) and a German translation (see Appendix I1) of the Multiple Stimulus Types Ambiguity Tolerance Scale–2 (MSTAT–2; McLain, 2009).

In part (4) an IAT (either valence or arousal version) with seven blocks was done. The response keys were E (left) and I (right). Blocks 1, 2, 3 and 6 consisted of 20 trials, blocks 4, 5 and 7 had 40 trials each. The a priori defined “congruent” pairings (used in block 3 and 4) of ambiguity with negative (activated) and clearness with positive (deactivated) always preceded the “incongruent” pairings (block 6 and 7). We used only a single sequence of congruency, as counterbalancing the order of congruency would lead to an increase of variance in the *D* measures which does not reflect variance of interindividual differences of strength of associations, but method-specific variance. For further discussion of counterbalancing or not counterbalancing the sequence of congruency especially in research interested in interindividual differences see Egloff and Schmukle (2002). The initial position (they switched positions between block 4 and 5) of target concepts ‘mehrdeutig’ [‘ambiguous’] (left position) and ‘eindeutig’ [‘clearness’] (right position) was fixed across participants. They were presented in green color. Attribute concepts were presented in white. An incorrect response was visualized by a red X which remained on the screen for 300 ms. The ITI was 250 ms.

The last part (5) was subdivided in two parts. In the first part, participants in the valence (arousal) condition had to rate how negative or positive (activated or deactivated) they felt while reading 20 (10 ambiguous and 10 unambiguous) sentences in random order (see Appendix B10). To this end, a 9-point rating scale was used. On every second point the corresponding (valence or arousal) manikin from the self-assessment manikin (see Bradley & Lang, 1994) was displayed. No further verbal anchor was presented during the rating phase. The 10 ambiguous and 10 unambiguous sentences were predominantly the same except for a lexical (un)ambiguous word. The same matched 10 lexical ambiguous and 10 unambiguous words as in the priming paradigm were used. In the second part, a sentence completion task was administered. The task was to choose one of three possible sentence endings of an ambiguous sentence (see Appendix B11). One ending did not disambiguate the initial sentence; the two other endings each highlight one of two possible meanings. Finally, participants could provide information (not saved together with previous data) for the lottery or for course credit.

3.4.3 Results

We used a multilevel model with random intercepts for the 20 prime words, 20 target words and participants. Calculations of models were done by the *lmer* function from the *lme4*-package (Bates et al., 2015) in RStudio (R Core Team, 2017). In the fixed effects model, we regressed the logarithmized reaction times in the priming paradigm on prime type × target type × IAT *D*-value and all subordinate two-way interactions and the main effects. In order to bind error variance, we included further interactions: In the EP (AP) paradigm we included a two-way interaction of the valence (arousal) of the prime with the target type. For the EP paradigm this two-way interaction is the standard EP effect. We tried to use only primes with neutral valence but of course there is some in valence of the primes around the (neutral) mean. In order to control for this source of variance in the latencies, we also included the valence of the prime and its interaction with target valence. For the EP paradigm, a significant standard EP effect can be taken as support for validity of data.

Taken into account that on the one hand the EP effect is driven by response competition between the automatically activated response of the prime and the selection of the correct response for the target (Klauer, 1997; Klauer et al., 2005; Klinger et al., 2000; Voss et al., 2013) and that there are trial-by trial effects in the EP paradigm (Frings & Wentura, 2008), we decided to include the target type of the previous trial (target type N-1) and its two-way interaction with target type of the current trial to control for this potential source of variance.

All trials in the priming paradigms with reaction times below 300 ms (EP: 0.21 %; AP: 3.18 %) and above 3000 ms (EP: 0.09 %; AP: 0.65 %) and all false classifications (EP: 2.75 %; AP: 15.44 %) were excluded from the analysis. This meant a combined total exclusion of 2.95 % (17.52 %) of EP (AP) trials. There were 19,944 (11,085) remaining valid trials for the EP (AP) paradigm.

We used the improved algorithm (Greenwald et al., 2003, Table 4) to calculate the *D* score of the IAT. Because our IAT did not have a built-in penalty (forcing the participant to select the correct response key after making a mistake / wrong responses) we used the recommended 600 ms penalty for false classifications. As in study 1, a positive value indicated that ambiguity was associated more strongly with negativity (high arousal) and clearness more with positivity (low arousal) than ambiguity with positivity (low arousal) and clearness with negativity (high arousal). According to the procedure for the calculation of the *D* score, we excluded 3 (1) participants of the valence (arousal) IAT data (those who responded in less than 300 ms in more than 10 percent of trials). Therefore, a *D* score was calculated for 164 (111) participants.

The mean *D* score was -0.03 (*SD* = 0.38) for the valence-IAT and 0.41 (*SD* = 0.47) for the arousal-IAT. The odd even reliability of the valence-IAT was .73 and .83 for the arousal-IAT. Figure 9 shows mean latencies (after data trimming according to the improved algorithm for the *D* score) and standard errors for congruent and incongruent blocks for valence and arousal IAT.

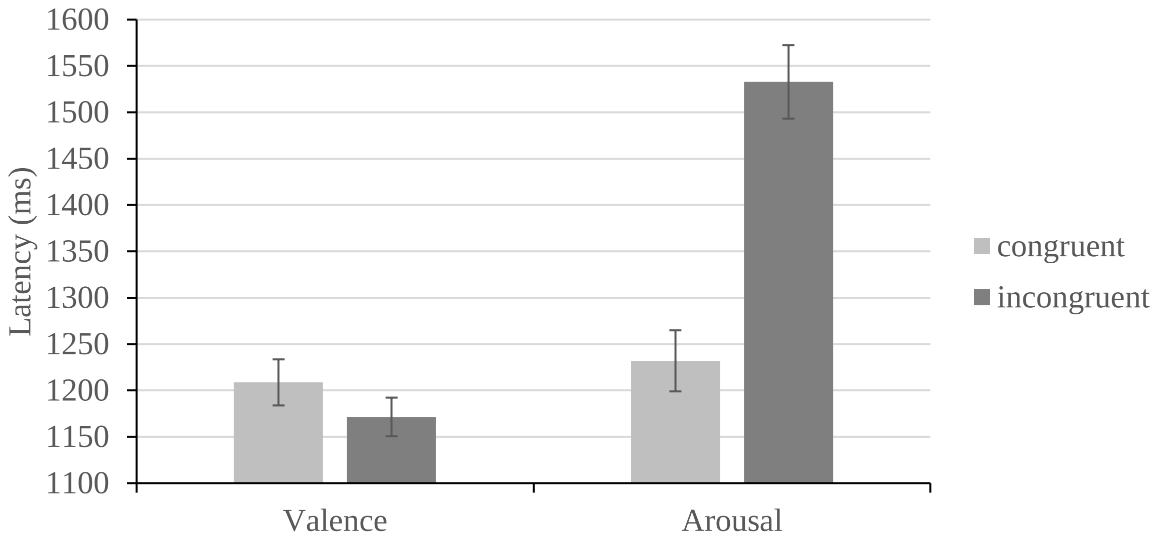


Figure 9. Mean Latencies and standard errors for congruent (ambiguous + negative/activated) and unambiguous + positive/deactivated) and incongruent (ambiguous + positive/deactivated and unambiguous + negative/activated) blocks of the IAT.

All *p* values reported for fixed effects of mixed-effect models were estimated by using the Satterthwaite approximation, which does not seem to be anti-conservative in terms of Type 1 error especially for restricted maximum likelihood models (Luke, 2017).

Valence. The specified mixed-effect model was based on 18,656 observations from 164 participants. They (118 females, 45 males, one missing gender information) had a mean age of 28.10 years (*SD* = 9.85 years). Random effect estimation for variance (*SD*) for the intercepts of participants was 0. 0180 (0.1343), for the target stimuli it was 0.0009 (0.0302), and 0.000 (0.0000) for the prime stimuli. Fixed effects of the model with an uncentered *D* score can be seen in Table 18. As the uncentered mean *D* score is very close to zero (*M* = -0.03), we will report only the uncentered results.

Table 18

Fixed effects of specified mixed-effect model (study 2 - evaluative priming).

	b	*SE b*	*df*	*t*	*p*
Intercept	6.4818	0.0126	123	514.75	< .001
Target type N-1	0.0101	0.0014	18466	7.34	< .001
Target type	0.0191	0.0069	18	2.77	.01
Prime valence	0.0002	0.0018	18466	0.14	.89
Prime type	0.0031	0.0014	18466	2.26	.02
D	0.0224	0.0277	162	0.81	.42
Target type N-1 × Target type	0.0056	0.0014	18470	4.02	< .001
Target type × Prime valence	0.0019	0.0018	18466	1.07	.29
Target type × Prime type	-0.0030	0.0014	18466	-2.17	.03
Prime type × *D*	0.0091	0.0036	18466	2.52	.01
Target type × *D*	-0.0011	0.0036	18464	-0.29	.77
Target type × Prime type × *D*	-0.0074	0.0036	18465	-2.06	.04

Note. Effect coding was used for the dichotomous variables: Target type and Target type N-1 (-1 = positive, 1 = negative), Prime type (-1 = unambiguous, 1 = ambiguous). Prime valence was centered by the mean. The *D* score was uncentered (the sample mean *D* score was -0.03).

The target type N-1 effect showed longer latencies given negative (vs. positive targets). This was true for the target type effect, too. The prime type effect revealed longer latencies in the case of ambiguous (vs. clear) primes. The included interaction of target type N-1 × target type was significant. This interaction revealed shorter latencies when the previous and the current target type had the same valence. Unexpectedly, the target type by prime type interaction was significant (see Figure 10). This means that for a *D* score of zero the influence of prime type differs across the positive and negative targets. The interaction plot can be seen in Figure 10. As can be seen there is no influence of prime type for negative targets (*b* = 0.0001, *SE b* = 0.0020, *t* = 0.06, *p* = .95) but a simple main effect of prime type for positive Targets (*b* = 0.0061, *SE b* = 0.0020, *t* = 3.14, *p* = .002). We will return to this result in the discussion.

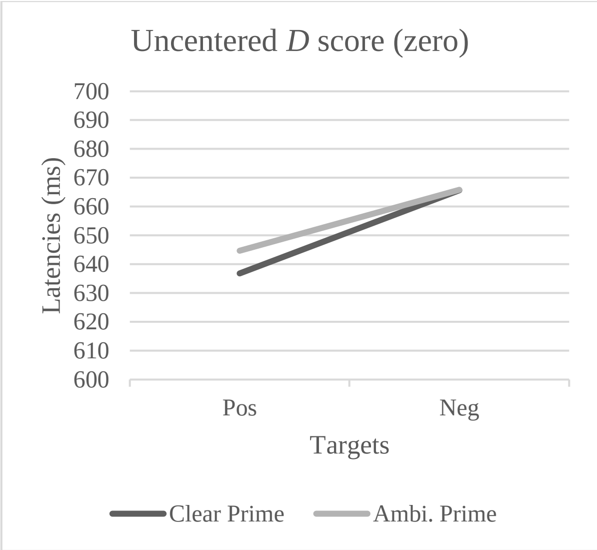


Figure 10. Predicted latencies as a function of prime type and target type for an uncentered *D* score of zero.

Also unexpectedly, the interaction of prime type and *D* was significant. This interaction effect revealed longer latencies in the case of ambiguous (vs. clear) primes for higher values of *D*. However, these main effects and two-way interaction must be interpreted in light of the significant three-way interaction of prime type, target type, and *D* score. The moderation of the prime type × target type interaction by the *D* value of the IAT can be seen in Figure 11.

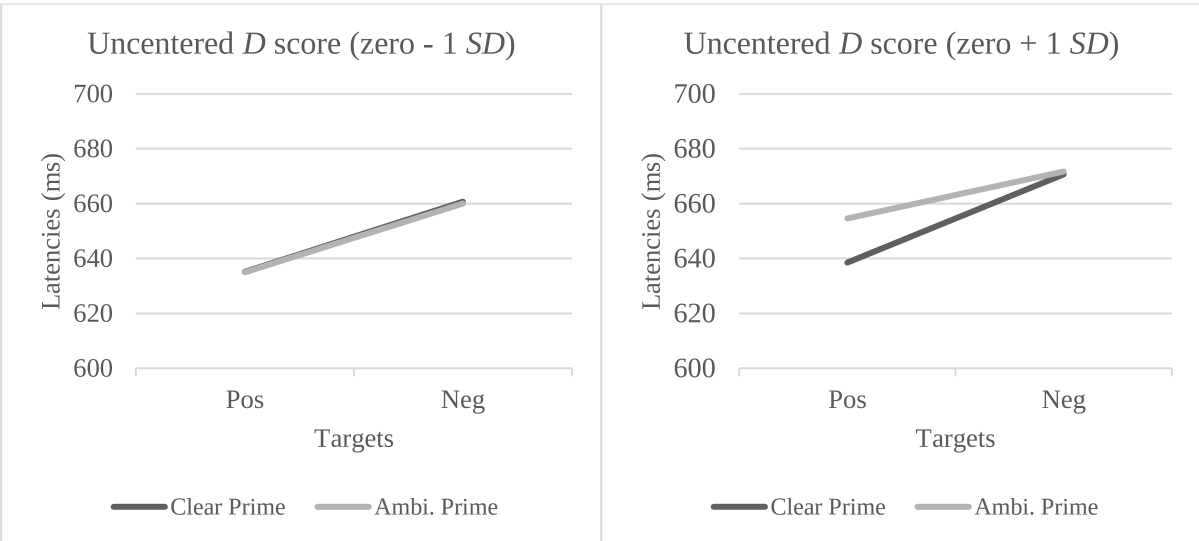


Figure 11. Predicted latencies as a function of prime type and target type for positive and negative *D* score. Left (right) plot shows the predicted values for -1 (+1) *SD* (0.38) below (above) the uncentered *D* score. Therefore, the left (right) plot shows predicted latencies for participants who like ambiguity more than clearness (clearness more than ambiguity).

Contrary to expectations, we did not find a significant interaction of prime type and target type for a zero *D* score – 1 *SD* (*b* = -0.0002, *SE b* = 0.0019, *t* = -0.09, *p* = .93). That means that for participants for which the IAT indicates that they like ambiguity more than clearness the influence of prime type is not differ between positive and negative targets. In contrast, and in line with the predictions, for a zero *D* score + 1 *SD* (which means stronger liking of clearness than ambiguity or stronger disliking of ambiguity than clearness) the interaction of prime type × target type was significant (*b* = -0.0058, *SE b* = 0.0020, *t* = -2.86, *p* = .004). The simple main effect of prime type given positive targets for a zero *D* score + 1 *SD* (*b* = 0.0125, *SE b* = 0.0028, *t* = 4.31, *p* < .001) revealed longer latencies given ambiguous (vs. clear) primes. This was as predicted. Participants for which the IAT *D* score indicated stronger disliking of ambiguity (or less liking of ambiguity) compared with clearness should have stronger negative (or weaker positive) evaluation for ambiguous compared to clear primes which leads to an evaluative conflict with the positive targets. The reversed pattern (longer latencies given clear (vs. ambiguous) primes was predicted for negative targets. As can be seen in Figure 11, this expected simple main effect is not significant (*b* = 0.0008, *SE b* = 0.0029, *t* = 0.27, *p* = .79).

Arousal. 10,797 observations from 111 participants (85 females, 26 males) built the basis for the mixed-effect model. They had a mean age of 27.64 years (*SD* = 11.10 years). Random effect estimation for variance (*SD*) for the intercepts of participants was 0.0462 (0.2150), for the target stimuli it was 0.0024 (0.0485), and 0.0000 (0.0000) for the prime stimuli. Fixed effects of the model can be seen in Table 19.

Table 19

Fixed effects of specified mixed-effect model (study 2 - arousal priming).

	b	*SE b*	*df*	*t*	*p*
Intercept	6.6048	0.0292	124	226.03	< .001
Target type N-1	-0.0028	0.0023	10659	-1.24	.22
Target type	-0.0230	0.0113	19	-2.04	.06
Prime arousal	0.0010	0.0034	10658	0.29	.77
Prime type	-0.0049	0.0031	10661	-1.60	.11
D	-0.0737	0.0435	108	-1.69	.09
Target type N-1 × Target type	0.0002	0.0023	10660	0.10	.92
Target type × Prime arousal	0.0023	0.0034	10659	0.67	.50
Target type × Prime type	-0.0021	0.0031	10661	-0.67	.50
Prime type × *D*	0.0064	0.0049	10659	1.29	.20
Target type × *D*	0.0030	0.0050	10664	0.61	.54
Target type × Prime type × *D*	0.0030	0.0049	10659	0.62	.54

Note. Effect coding was used for the dichotomous variables: Target type and Target type N-1 (-1 = low arousal, 1 = high arousal), Prime type (-1 = unambiguous, 1 = ambiguous). Prime arousal was centered.

As the expected three-way interaction was not significant, we do not report predicted values for one *SD* below and above the mean *D* score.

3.4.4 Discussion

In this study we investigated whether lexically ambiguous and unambiguous words trigger an automatic activation of valence and arousal assessed with two different sequential priming paradigms, and especially, whether there is evidence for an interindividually different automatic activation of valence and arousal. Therefore, we additionally measured the strength of associations of the concepts of ambiguity and clearness with positive (low) or negative (high) valence (arousal), respectively, via the IAT. Our results provide evidence for interindividually different automatic valence (but not arousal) activation as the prime type × target type interaction (EP paradigm) was, as predicted, moderated by the IAT *D* value.

Obviously, this significant three-way interaction only partly reflects our hypothesis. Firstly, a significant interaction of prime type × target type was found only for participants for which the IAT *D* value indicated a relative liking of ambiguity over clearness, and secondly, this interaction was only driven by the influence of the primes for positive targets and not for negative targets. The latter finding is in line with results reported by Unkelbach et al. (2008). They showed that the EP effect is driven more strongly by the congruent pairing of positive primes with positive targets compared to the pairing of negative primes with negative targets. They explained this with the density hypothesis which states that positive information is more similar to each other than negative information. Therefore, positive primes facilitate reactions to positive targets more strongly compared to the facilitation effect of negative primes for negative targets. Another possible explanation comes from results shown in our data: as is common in the EP paradigm (Dijksterhuis & Aarts, 2003; Unkelbach et al., 2008), negative targets were classified more slowly than positive targets. It could be speculated that the already weak valence activation derived from the prime (weak as the prime words were all close to neutral in valence, so valence should only result from individuals' attitudes towards ambiguity, which can be assumed not to be a strong attitude object) loses its impact on the target because of these longer response times for negative targets. Potential reasons for this could be, for example, functional segmentation of the prime from the target (see, Fiedler et al., 2011). However, the process behind the finding of stronger priming for positive targets is beyond the scope of this study.

It seems more important to explain, why prime type × target type interaction was not significant for participants for whom the IAT score indicated relative liking of ambiguity over clearness (as compared to the prime type × target type interaction for participants for whom the IAT score indicated a relative disliking of ambiguity over clearness). Although ambiguity occurs in daily life, it can be suspected that the default state of information which we are confronted with is clearness: most what we see, listen, smell, touch, and taste is probably not ambiguous. Therefore, a strong evaluative reaction to clearness and a strong association with positive or negative valence seems unlikely. In the reverse conclusion, although the IAT score usually reflects a relative liking/disliking of two concepts, in this case it could potentially be interpreted in an absolute manner. Under the assumption of a neutral attitude towards clearness, a negative (positive) ambiguity-IAT score would indicate an association of ambiguity with positivity (negativity) rather than reflecting the liking/disliking of ambiguity in relation to clearness. Together with the assumption that negative events elicit a stronger initial emotional and cognitive response than positive events (Taylor, 1991), we would expect a smaller influence of the positively associated ambiguous primes compared to negatively associated ambiguous primes in the EP paradigm. Especially so, if we further assume that the evaluative reaction to clear primes does not strongly differ interindividually. This differential reaction to ambiguity could explain why we only found an interaction of prime type × target type for persons for whom the IAT reflected a stronger association of ambiguity with negative valence (and clearness with positive valence) than clearness with negative valence (and ambiguity with positive valence).

An alternative explanation could be the influence of third variables that affect both the IAT and the EP paradigm in such a way that we found the observed relationship between these two indirect measures. As mentioned, by using the improved algorithm, method-specific variance like task-switching costs (Mierke & Klauer, 2003), task-switching abilities (Back, Schmukle, & Egloff, 2005) and cognitive abilities (Wright & Meade, 2012) was reduced in the IAT score. However, it could still be that the ambiguity-IAT indirectly measures cognitive abilities. This could be the case if there was a relationship between the attitude towards ambiguity and cognitive abilities, possibly in a way that liking of ambiguity is linked to higher cognitive abilities. Indirect evidence for such a relationship comes from the known relationship of TA with need for cognition (Wolfradt, Oubaid, Straube, Bischoff, & Mischo, 1999) together with the known relationship of need for cognition with intelligence (Hill et al., 2013). However, studies assessing both TA and a measure of cognitive ability showed mixed results in terms of correlation between these two concepts (Jach & Smillie, 2019; Raphael, Moss, & Cross, 1978).

Nonetheless, if the ambiguity-IAT indirectly assesses cognitive abilities, this could offer an explanation for the non-existent prime type × target type interaction for participants for whom the IAT indicates a relative liking of ambiguity over clearness. If these participants additionally have higher cognitive capacities, they are more likely to be able to strategically control for the influence of the prime on the target. Indeed, several studies demonstrated that the EP paradigm is sensitive to top-down influences, generally speaking (Alexopoulos, Fiedler, & Freytag, 2012; Alexopoulos et al., 2017; Fiedler et al., 2011).

Interestingly, in contrast to study 1, the mean *D* score was very close to zero, which can be interpreted to mean that the sample in general had no preference for ambiguity over clearness or vice versa. However, in this study we found a significant prime type × target type interaction. This unexpected result raises questions. It means that the EP paradigm indicated an automatic activation of valence of the ambiguous or clear primes for those persons for whom an IAT *D* score reflected no preference for ambiguity over clearness or vice versa. This could be interpreted in a way that a zero IAT *D* score does not necessarily reflect equal liking (or disliking) of ambiguity and clearness and hence only the relative position of an IAT score within a given distribution of a sample can be interpreted in a meaningful way, not the absolute value. Indeed, the three-way interaction supported the idea that the relative positions of the IAT scores have some predictive power. This strengthens the view that the ambiguity-IAT scores contain construct related variance. However, considering the mean IAT score as valid in a sense that the sample in general had no preference for clearness over ambiguity, leads to the question whether the prime type × target type interaction (especially the faster responses for positive targets following a clear prime compared to an ambiguous one) can be resolved by taking potential differences in the speed of meaning activation for ambiguous and unambiguous words into account. Consider the theoretical case that activating the meanings of ambiguous words takes much more time than activating the meanings of unambiguous words. This consideration together with the fact that both ambiguous and unambiguous words had on average a slightly positive valence (see matching results) could lead to differential opportunities for evaluative priming of ambiguous versus unambiguous words. However, this should only be the case if the delay for meaning activation for ambiguous words was so large that these words have no or only limited impact on the evaluative decision of the target. There is evidence from behavioral data that the speed of meaning activation for isolated words does not differ between ambiguous and unambiguous words (Pexman, Hino, & Lupker, 2004). However, a study using event-related brain potentials by means of magnetoencephalography (Beretta et al., 2005) found a difference in the M350 (assumed to reflect the initial lexical access) for homonyms (354 ms) and non-homonyms (336 ms). As there is evidence for no or only small differences in the speed of meaning activation between ambiguous and unambiguous words, we deem this potential interpretation of the prime type × target type interaction unlikely.

Another possible explanation for this result comes from the view that the IAT score is more likely to reflect the attitude towards the target concepts in relation to each other on a conceptual level (and that a zero IAT score does indeed reflect no preference for one target concept over the other) and the EP paradigm is more likely to reflect the attitude on a stimulus driven level (see 2.3.2). Moreover, as indirect measures are affected by both automatic and controlled processes (J. W. Sherman, 2008), there could be differences in the extent to which these two processes play a role in the IAT and the EP paradigm.. These assumed differences (measurement level and process “purity”) between the two indirect measures could be relevant, too, in order to understand why a zero IAT effect does not necessarily correspond to a nonsignificant prime type × target type interaction in the EP paradigm.

Because of difficulties in interpreting an IAT effect, the question arises whether there is a way of answering the question of interindividual differences in automatic evaluations of ambiguity without the IAT. A possible strategy would be to use the known group approach instead of relying on the measured attitude and using this information as a moderator. Although there are specific groups having low or high TA (see, e.g., Furnham & Ribchester, 1995), these differences were measured via direct measures of TA, which were characterized by a broad unspecific conceptualization of ambiguity (see 2.3.1). There were several reasons to dismiss the idea of using known groups instead of measured attitude via the IAT. On the one hand, the validity of the measures of TA used in previous research is questionable, which means the groups might not differ in our conceptualization of attitude towards ambiguity. On the other hand, there is less variance in a dichotomous variable, reducing power. Therefore, a known group approach does not seem superior for testing the assumption of interindividual differences in unintentional activations of evaluations of ambiguity. As mentioned, knowing about the low correlation of sequential priming measures with direct measures (Cameron et al., 2012), direct measures seem not to be an option for assessing the attitude towards ambiguity and using this information as a moderator, either.

Although there is strong evidence (see section 2.2) for multiple meaning activation for ambiguous words presented in isolation (especially for ambiguous words with balanced meanings) we included a learning paradigm in this study in which the participants were supposed to learn which word was ambiguous and which word was unambiguous prior to the EP paradigm. We did this in order to pre-activate both meanings of the ambiguous words. This was supposed to guarantee that all participants “saw” the ambiguity in the EP paradigm, in order to boost potential priming effects related to the factor ambiguousness. Through this, the external validity of the results is impaired in a sense that we cannot draw conclusions about interindividual differences in automatic activation of evaluations of ambiguous and clear words for which the multiple meanings for the ambiguous words were not pre-activated. Therefore, we cannot say if the mental state of having two meanings vs. one meaning in mind triggers an (interindividually different) automatic activation of evaluation as the learning prior to the EP paradigm could lead to an association of each word with the concepts of ambiguity or clearness, which is not the same as having one or two meanings in mind. Therefore, in study 3, we introduced a condition without a learning phase prior to the EP paradigm. With this additional condition, we aimed at drawing conclusions about the (interindividually different) automatic activation of evaluations of the automatic activated meaning(s) derived from perceiving ambiguous versus clear words (in case of successful replication of the three-way-interaction). Nonetheless, this study offered first evidence for interindividual differences in automatic activation of evaluations of ambiguous and unambiguous information.

We will not discuss the nonsignificant three-way interaction in the arousal condition in detail as on the one hand, to our knowledge only one published study applied the AP paradigm (but including only stimuli of positive valence) and found no evidence for arousal priming (Hinojosa, Carretié, Méndez-Bértolo, Míguez, & Pozo, 2009) and on the other hand, the high error rate and slow responses (M = 735 ms) in the AP paradigm can be interpreted to suggest that this task was highly demanding for participants and hence potentially reduced the impact of the primes. On a side note, we considered the AP paradigm falling into the class of categorical priming for which (if the irrelevant primes were automatically classified in terms of the relevant target categories) response competition can be expected (see, Voss et al., 2013). Therefore, one possibility for the result is that no automatic arousal-categorization of the primes occurred and thus no congruency effect (moderated by the IAT) was evident.

3.5 Study 3

This study represents a conceptual replication of the EP paradigm and the valence-IAT conducted in study 2. Therefore, we tried to replicate the partial evidence for automatic activation of valence of lexically ambiguous and unambiguous (clear) words depending on the measured associations (via an IAT) of the concepts of ambiguity and clearness with positive or negative valence, respectively. Based on the strong evidence for multiple meaning activation for isolated presented (especially balanced) ambiguous words (see section 2.2), we expected that a learning paradigm prior to the EP paradigm is not necessary in order to find evidence for interindividual differences (corresponding to the IAT result) in the automatic activation of evaluations of ambiguous and unambiguous prime words. Therefore, we added a condition without a learning phase prior to the EP paradigm. In order to increase power, we increased the number of trials and stimuli.

3.5.1 Hypothesis

Again, we expected that the prime × target interaction was moderated by the IAT *D* value for both the condition with a learning phase and for the condition without a learning phase (cf. hypothesis of study 2 and Figure 8).

3.5.2 Method

Participants. Data were collected for one month from March 13 to April 13, 2018. All participants were recruited via the mailing list of the University of Tübingen. We only included data from participants who finished the study completely. From these 452 participants, we excluded 9 participants who declared to speak German not on native speaker level. Nobody declared to participate via smartphone. The remaining 443 participants (347 females, 91 males, 5 missing gender information) had a mean age of 25.81 years (SD = 8.38 years). 241 participants took part in the learning phase prior to the EP paradigm and 202 started the EP paradigm without the learning phase. The first 150 participants received a voucher over the amount of 5 Euro (accepted by many shops in Tübingen). The rest could take part in a lottery for one of ten 20 Euro for a chocolate shop located in Tübingen.

Design. Participants were randomly assigned to the learning phase condition (present vs. absent). A repeated measure design was used for the two factors prime type (ambiguous vs. unambiguous) and target type (positive vs. negative).

Materials. Stimulus words were selected from the NAUG database (see Appendix A3 - A5) and matching was done with the program Match (van Casteren & Davis, 2007). As in study 2, we preselected ambiguous and unambiguous words prior to matching. First, we excluded all ambiguous words for which the frequencies of the first and second meaning differed strongly from an equal distribution ($\chi^2 > 30$). Secondly, we excluded all ambiguous words for which there were more than 20 percent of associations which did not match the first two meanings. Then we eliminated all ambivalent ambiguous words for which the valence ratings of the first and second meaning differed strongly (difference of 2.5 or higher) from each other, but only for those words for which the mean valence rating of one meaning was below the scale mean of 5 and the other mean valence rating was above the scale mean. From the unambiguous words all strongly positive (mean rating above 7) and all strongly negative (mean rating below 3) words were used for matching of target stimuli. The remaining unambiguous stimuli were used for matching with the ambiguous stimuli. In order to have more variance in valence of both ambiguous and unambiguous primes, we divided the ambiguous words into three categories: more negative (< 4.5), neutral (> 4.5 and < 5.5) and more positive (> 5.5). An acceptable match was found for an output size of 6 items pairs for the positive and negative sets and an output size of 8 item pairs for the neutral set. Matching dimensions (with higher emphasis on the valence dimension) were: word frequency, number of letters, valence, arousal and abstractness ratings of the word. For the last three dimensions the mean of the mean ratings of both meanings were used. The 20 unambiguous and 20 ambiguous words (see Appendix E1; including *M* and *SD*) did not differ in frequency, $t(38) = 0.30$, $p = .77$, number of letters, $t(38) = 0.12$, $p = .90$, valence of the word $t(38) = -0.54$, $p = .59$, mean valence of meanings $t(38) = 0.26$, $p = .80$, arousal of the word $t(38) = 0.08$, $p = .93$, mean arousal of meanings $t(38) = -0.20$, $p = .84$, abstractness of the word $t(38) = 0.83$, $p = .41$, or the mean abstractness of meanings $t(38) = 0.66$, $p = .51$. The matched 10 positive and 10 negative words used as targets (see Appendix E2; including *M* and *SD*) did not differ in frequency, $t(18) = -0.37$, $p = .71$, number of letters, $t(18) = 0.16$, $p = .87$, arousal of the word $t(18) = -0.72$, $p = .48$, abstractness of the word $t(18) = 0.01$, $p = .99$ but as intended in valence $t(18) = 29.92$, $p < .001$.

For the IAT we used the 5 best matching pairs from the stimuli from the sequential priming paradigm. The 5 ambiguous and 5 unambiguous words (see Appendix E3; including *M* and *SD*) did not differ in frequency, $t(8) = -0.45$, $p = 66$, number of letters, $t(8) = 0.41$, $p =$

.69, valence of the word $t(8) = -1.54$, $p = .16$, mean valence of meanings $t(8) = 0.08$, $p = .94$, arousal of the word $t(8) = 0.45$, $p = .67$, mean arousal of meanings $t(8) = 0.23$, $p = .82$, abstractness of the word $t(8) = 1.45$, $p = .18$, and abstractness of meanings $t(8) = 1.55$, $p = .16$. The 5 positive and 5 negative words (see Appendix E4; including *M* and *SD*) did not differ in frequency, $t(8) = -0.53$, $p = .61$, number of letters, $t(8) = -0.63$, $p = .54$, arousal of the word $t(8) = -0.35$, $p = .73$, abstractness of the word $t(8) = -0.36$, $p = .73$ but as intended in valence $t(8) = 32.22$, $p < .001$.

Procedure. The online study used the JavaScript library jsPsych (de Leeuw, 2015). It ran in full-screen mode with black font on white background on the PC or laptop of the participants. The learning phase differed from study such that this time only isolated words were presented in the center of the screen and the task for the participants was to decide if the word has one or two meanings by pressing the corresponding key. Correct responses were marked by the letter string “Richtig” [“Correct”] for 700 ms. Incorrect responses were marked by the letter string “Falsch” [“Wrong”] and the two meanings in case of wrong classification of an ambiguous word or a notion that the word has only one meaning in case of wrong classification of an unambiguous word was presented for 4500 ms. The procedural change compared to study 2 in which the same two sentences were used for both the ambiguous and the unambiguous words for the learning phase was necessary because meaningful sentences for the 20 ambiguous and 20 unambiguous words could not be constructed and would take too much time for the participants to deal with.

After the learning phase the participants conducted the EP paradigm. 40 training trials with feedback (300 ms) after each trial were followed by the three test blocks with 80 trials each without feedback. As in study 2, a trial consisted of a fixation cross (500 ms), a blank screen (10 ms), the prime (200 ms), a blank screen (100 ms) and the target which was displayed until a response was given. This time response keys were E for negative words and I for positive words. The ITI was 810 ms. Between blocks participants could take a break. Again, participants were told that they should answer as fast as possible.

The following IAT differed only in two aspects from the one used in study 2. Firstly, in order to continue after an incorrect response, the participants had to press the correct response key (built-in penalty). Secondly, this time attribute concepts and target concepts were all presented in black font.

After the IAT, an experiment designed to investigate a potential mood change while conducting an (un)ambiguous task depending on the attitude towards ambiguity was done. Because this is a separate research question not relevant to test the hypothesis described, we will not comment on this part here. Subsequently the explicit measures were administered. This part consisted of the Uncertainty Tolerance Scale (Dalbert, 1999), 14 items derived from the confirmatory factor analysis of study 2 and 7 new added items (see Appendix B12) and a German translation (see Appendix I) of the Multiple Stimulus Types Ambiguity Tolerance Scale–2 (MSTAT–2; McLain, 2009). At the end, participants could provide information (not saved together with previous data) for the lottery or for course credit.

3.5.3 Results

We applied a multilevel model with identical random effects, like in study 2. The fixed effects model was extended (compared to study 2) by the factor learning phase with the main effect and all interactions but otherwise identical. Like in study 2, we excluded all trials in the priming paradigms with reaction times below 300 ms (0.24 %) and above 3000 ms (0.14 %) and all false classifications (4.20 %) from the analysis. In total, we excluded 4.46 % of EP paradigm trials. This resulted in 101,573 remaining valid trials for the EP paradigm.

We again used the improved algorithm to calculate the *D* score of the IAT, but this time RT of false responses included a built-in penalty (forced correct key press after wrong classification). The time to respond to the correct response key after making a mistake took a median of 588 ms (mean 911 ms) which is compatible to the 600 ms penalty used in study 2. As in study 2, we calculated the *D* score so that a positive value indicated that ambiguity was associated more strongly with negativity and clearness more with positivity than ambiguity with positivity and clearness with negativity. The IAT data of 5 participants (those who responded in less than 300 ms in more than 10 percent of trials) were excluded. Therefore, a *D* score was calculated for 438 participants.

The mean *D* score for all participants was 0.25 (*SD* = 0.36). For the participants with learning phase the mean *D* score was lower 0.19 (*SD* = 0.34) than for the participants without learning phase 0.31 (*SD* = 0.36), $t(436) = -3.72$, $p < .001$. The odd even reliability for the *D* score was .64 (.68) for participants with (without) leaning phase. The latencies in ms (derived from the application of data trimming according to the improved algorithm for the *D* score) for the congruent block were shorter (*M* = 1089, *SD* = 279.10) than for the incongruent block (*M* = 1253, *SD* = 343.87), $t(437) = 12.78$, $p < .001$. For the condition with learning phase the mean latency in ms for the congruent (incongruent) block was *M* = 1108 *SD* = 291.51 (*M* = 1232 *SD* = 341.72), the difference was significant with $t(237) = 7.51$, $p < .001$. For the condition without learning phase the mean latency in ms for the congruent (incongruent) block was *M* = 1067, *SD*

= 262.58 (*M* = 1279, *SD* = 345.67)), the difference was significant with $t(199) = 10.78$, p < .001.

The specified mixed-effect model was based on 99,169 observations from 438 participants. They (345 females, 88 males, 5 missing gender information) had a mean age of 25.79 years (*SD* = 8.40 years). Random effect estimation for variance (*SD*) for the intercepts of participants was 0.0198 (0.1409), 0.0015 (0.0383) for the target stimuli and 0.0000 (0.0023) for the prime stimuli. Fixed effects of the model using the uncentered *D* score can be seen in Table 20.

Table 20

Fixed effects of mixed-effect model with uncentered D score (study 3).

	b	*SE b*	*df*	*t*	*p*
Intercept	6.4209	0.0120	67	532.94	< .001
Target type N-1	0.0106	0.0006	98696	16.67	< .001
Target type	0.0227	0.0086	18	2.64	.02
Prime valence	-0.0010	0.0007	37	-1.45	.16
Prime type	0.0009	0.0009	76	1.04	.30
D	-0.0225	0.0193	434	-1.16	.25
Learning Phase (LP)	0.0093	0.0085	434	1.10	.27
Target type N-1 × Target type	-0.0165	0.0006	98704	-25.79	< .001
Target type × Prime valence	0.0043	0.0006	98666	6.93	< .001
Target type × Prime type	0.0004	0.0008	98664	0.51	.61
Prime type × *D*	-0.0003	0.0018	98664	-0.16	.87
Target type × *D*	-0.0026	0.0018	98663	-1.41	.16
Prime type × LP	0.0008	0.0008	98664	1.04	.30
Target type × LP	0.0008	0.0008	98663	0.97	.33
D × LP	-0.0219	0.0193	434	-1.13	.26
Prime type × Target type × *D*	-0.0019	0.0018	98662	-1.03	.30
Prime type × Target type × LP	-0.0016	0.0008	98663	-2.08	.04
Prime type × *D* × LP	-0.0001	0.0018	98664	-0.08	.94
Target type × *D* × LP	-0.0046	0.0018	98664	-2.52	.01
Prime type × Target type × *D* × LP	0.0034	0.0018	98663	1.89	.06

Note. Effect coding was used for the dichotomous variables: Target type and Target type N-1 (-1 = positive, 1 = negative), Prime type (-1 = unambiguous, 1 = ambiguous), Learning phase (-1 = without, 1 = with). Prime valence was centered by the mean. The *D* score is uncentered.

As can be seen, there was a main effect of target type. Negative (vs. positive) targets elicited longer latencies. The same was true if the previous target was negative (vs. positive), see target type N-1 effect. Like the previous study, the included interaction of target type N-1 × target type was significant and revealed shorter latencies when the previous and the current target type had the same valence. The significant target type by prime valence interaction showed the standard EP effect: latencies were shorter if the valence of the prime was congruent with the valence of the target. Different from study 2, the prime type by target type interaction was not significant. The predicted three-way interaction of target type, prime type, and *D* was also not significant. Instead, there was a marginally significant four-way interaction, which can be seen in Figure 12.

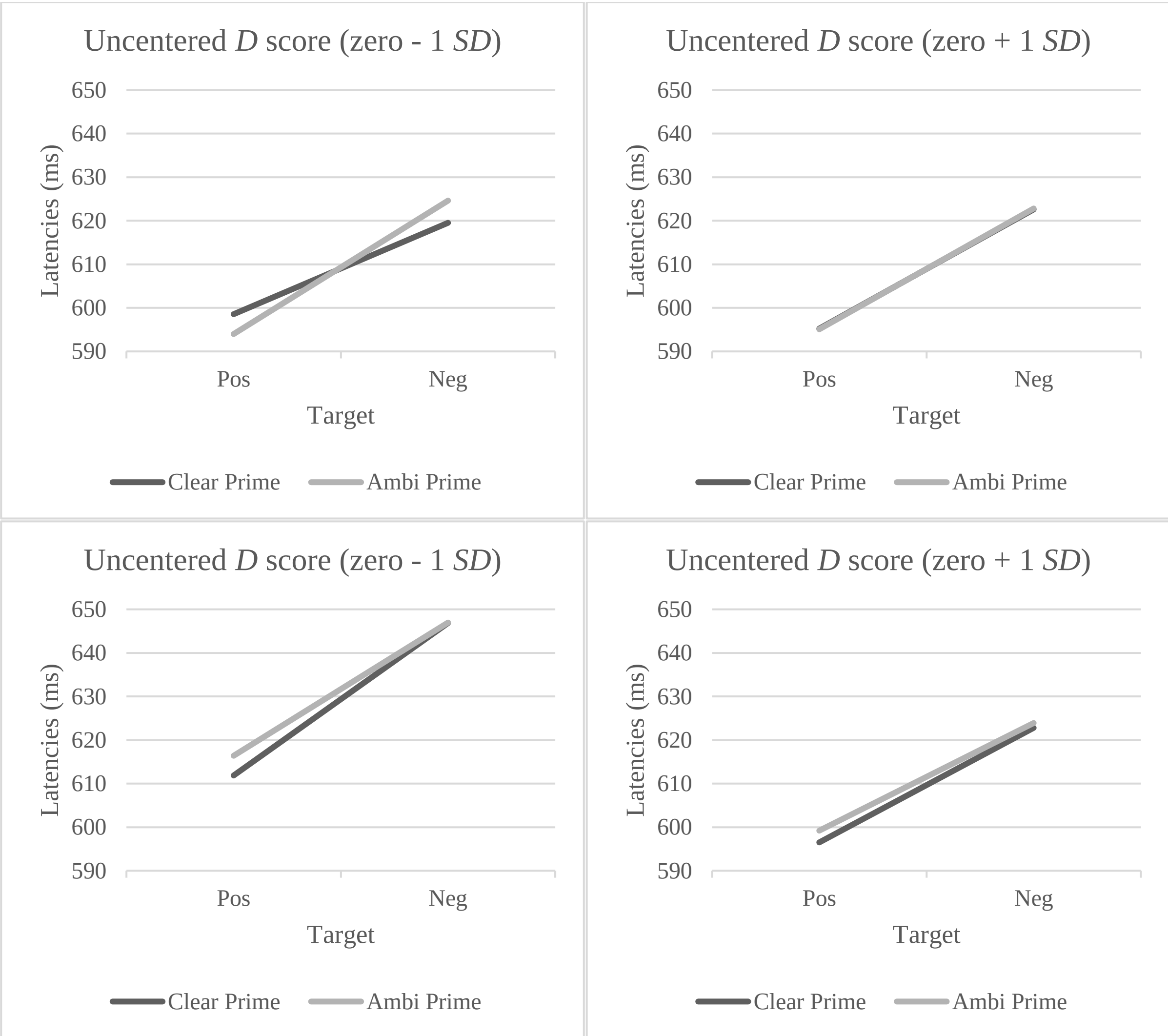


Figure 12. Interaction of prime type, target type, IAT *D* scores and learning phase. Upper (lower) two plots show predicted values for absent (present) learning phase. Left (Right) plots shows the predicted values for the -1 (+1) *SD* (0.36) below (above) the uncentered *D* score of zero. Therefore, left (right) plots show the predicted latencies as a function of prime type and target type for participants for which the IAT indicates that they like ambiguity more than clearness (clearness more than ambiguity).

To further investigate the four-way interaction, we dummy coded the variable LP. Table 21 shows the fixed effects with dummy coded variable LP (without LP coded with 0 and with LP coded with 1).This analysis showed that the three-way interaction of prime type × target type × *D* was significant for the condition without LP (see Figure 12, upper panel;

corresponding estimates in Table 21). The reversed coding of LP showed that for the condition with LP the interaction of prime type × target type × *D* was not significant, *b* = 0.0015 (SE *b* = .0025), *t*(98663) = 0.62, p = .54 (see Figure 12, lower panel).

Table 21

Fixed effects of mixed-effect model with uncentered D score for LP dummy coded (study 3).

	b	*SE b*	*df*	*t*	*p*
Intercept	6.4117	0.0158	167	405.52	< .001
Target type N-1	0.0106	0.0006	98696	16.67	< .001
Target type	0.0219	0.0087	19	2.53	0.02
Prime valence	-0.0010	0.0007	37	-1.45	0.16
Prime type	0.0001	0.0013	370	0.06	0.95
D	-0.0006	0.0278	434	-0.02	0.98
Learning Phase (LP)	0.0186	0.0169	434	1.10	0.27
Target type N-1 × Target type	-0.0165	0.0006	98704	-25.79	< .001
Target type × Prime valence	0.0043	0.0006	98666	6.93	< .001
Target type × Prime type	0.0021	0.0012	98663	1.65	0.10
Prime type × *D*	-0.0002	0.0026	98664	-0.06	0.95
Target type × *D*	0.0020	0.0026	98664	0.77	0.44
Prime type × LP	0.0017	0.0016	98664	1.04	0.30
Target type × LP	0.0015	0.0016	98663	0.97	0.33
D × LP	-0.0438	0.0387	434	-1.13	0.26
Prime type × Target type × *D*	-0.0053	0.0026	98663	-2.03	0.04
Prime type × Target type × LP	-0.0033	0.0016	98663	-2.08	0.04
Prime type × *D* × LP	-0.0003	0.0036	98664	-0.08	0.94
Target type × *D* × LP	-0.0091	0.0036	98664	-2.52	0.01
Prime type × Target type × *D* × LP	0.0068	0.0036	98663	1.89	0.06

Note. Effect coding was used for the dichotomous variables: Target type and Target type N-1 (-1 = positive, 1 = negative), Prime type (-1 = unambiguous, 1 = ambiguous). Prime valence was centered by the mean. Learning phase (0 = without, 1 = with). Prime valence was centered by the mean. The *D* score was uncentered.

In order to further analyze the significant three-way interaction of prime type × target type × *D* for the condition without LP, we analyzed the effects of the prime type × target type interactions for one *SD* below and above zero. For a zero *D* score – 1 *SD* (upper left plot of Figure 12) the interaction was significant with *b* = 0.0039 (SE *b* = .0020), *t*(98663) = 1.98, p = .047. but not for a zero *D* score + 1 *SD* (upper right plot of Figure 12) with *b* = 0.0002 (SE *b* = .0009), *t*(98663) = 0.19, p = .85. The simple main effect of prime type for positive Targets for zero *D* score – 1 *SD* (*b* = -0.0038, *SE b* = 0.0028, *t* = -1.33, *p* = .18) and the simple main effect of prime type for negative Targets for a zero *D* score – 1 *SD* (*b* = 0.0041, *SE b* = 0.0028, *t* = 1.45, *p* = .15) were not significant.

As already mentioned, there is strong evidence that the IAT *D* score is driven by associative and non-associative processes. Therefore, a zero IAT *D* score does not necessarily reflect equal liking/disliking of ambiguity and clearness. Because of this problem when interpreting the absolute values of IAT *D* scores an alternative option is to use the information of the relative position of a given *D* within the distribution of *D* scores of the investigated sample. Therefore, we ran an additional analysis with mean centered *D* scores. As the mean of *D* scores for participants with and without LP were not identical, we used the group centered *D* scores. As the first analysis already showed that for the condition with LP no moderation of prime type × target type by *D* is existent, we only calculated effects for the condition without LP by dummy coding. The fixed effects using centered *D* scores can be found in Table 22.

Table 22

Fixed effects of mixed-effect model with LP-group mean centered D score (LP dummy coded).

	b	*SE b*	*df*	*t*	*p*
Intercept	6.4115	0.0132	93	486.61	< .001
Target type N-1	0.0106	0.0006	98696	16.67	< .001
Target type	0.0225	0.0086	18	2.61	0.02
Prime valence	-0.0010	0.0007	37	-1.45	0.16
Prime type	0.0000	0.0010	134	0.03	0.98
D	-0.0006	0.0278	434	-0.02	0.98
Learning Phase (LP)	0.0103	0.0136	434	0.76	0.45
Target type N-1 × Target type	-0.0165	0.0006	98704	-25.79	< .001
Target type × Prime valence	0.0043	0.0006	98666	6.93	< .001
Target type × Prime type	0.0004	0.0009	98664	0.42	0.68
Prime type × *D*	-0.0002	0.0026	98664	-0.06	0.95
Target type × *D*	0.0020	0.0026	98664	0.77	0.44
Prime type × LP	0.0016	0.0013	98663	1.28	0.20
Target type × LP	-0.0004	0.0013	98663	-0.34	0.73
D × LP	-0.0438	0.0387	434	-1.13	0.26
Prime type × Target type × *D*	-0.0053	0.0026	98663	-2.03	0.04
Prime type × Target type × LP	-0.0013	0.0013	98663	-1.06	0.29
Prime type × *D* × LP	-0.0003	0.0036	98664	-0.08	0.94
Target type × *D* × LP	-0.0091	0.0036	98664	-2.52	0.01
Prime type × Target type × *D* × LP	0.0068	0.0036	98663	1.89	0.06

Note. Effect coding was used for the dichotomous variables: Target type and Target type N-1 (-1 = positive, 1 = negative), Prime type (-1 = unambiguous, 1 = ambiguous). Prime valence was centered by the mean. Learning phase (0 = without, 1 = with). The *D* scores were separately centered by the sample mean for participants with and without LP.

Predicted values of latencies as a function of prime type and target type by high and low IAT *D* score (referring to the sample mean) can be found in Figure 13.

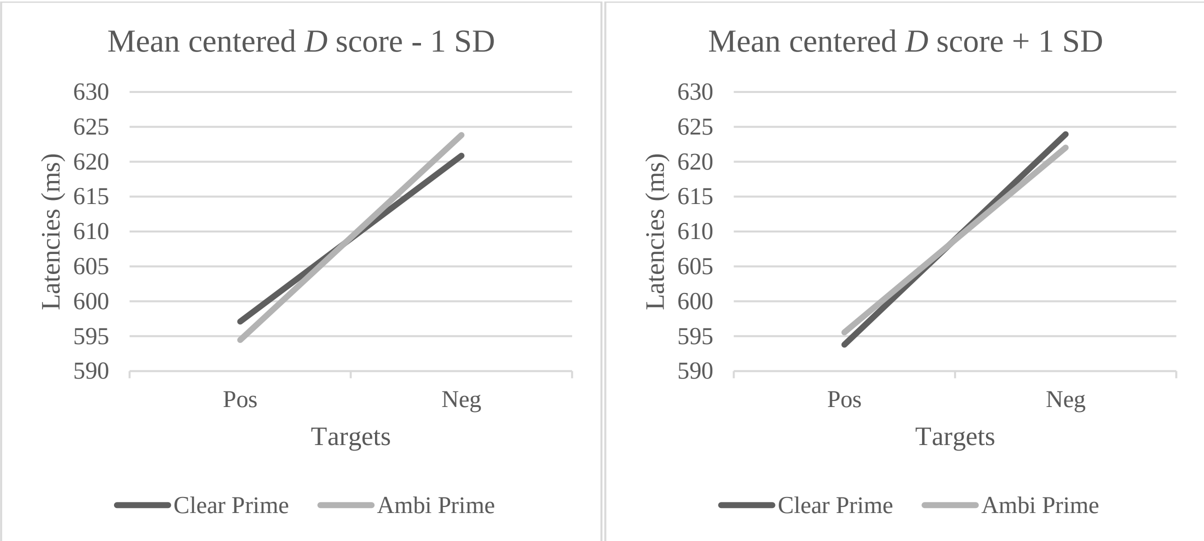


Figure 13. Moderation of latencies as a function of prime type and target type by IAT *D* scores. Left (Right) plot shows the predicted values for the -1 (+1) *SD* (0.36) below (above) the group mean centered *D* score (for the condition without LP: *M* = 0.31). Hence, when the sample average *D* score is considered as a neutral point, the left (right) plot shows the predicted values as a function of prime type and target type for participants liking ambiguity more than clearness (liking clearness more than ambiguity).

As can be seen in Figure 13, the pattern of predicted latencies as a function of prime type and target type for low and high IAT *D* score (when the sample mean is considered as neutral point) was similar to the expected pattern (cf. hypothesis of study 2 and Figure 8). However, the prime type × target type interaction for the mean centered *D* score – 1 *SD* (left plot of Figure 13) (b = 0.0023 (SE b = .0013), t(98663) = 1.72, p = .09) as well as the mean centered *D* score + 1 *SD* (right plot of Figure 13) with b = -0.0015 (SE b = .0013), t(98663) = -1.12, p = .26 were not significant.

Post-hoc analysis. As we used a subset of words from the EP paradigm as stimuli for the IAT, this could potentially lead to difficulties in interpreting the found three-way interaction of prime type, target type and IAT *D* score. If the IAT, different from assumptions, measures only random variation in valence over the lists of ambiguous and clear words, then the found effect could not be attributed to automatic activation of evaluations of the phenomenon ambiguity and clearness. However, this explanation could not be applied to trials in the EP paradigm for which neither the prime nor the target stimuli was part of the stimuli used in the IAT. Therefore, we ran an additional analysis with the information whether the stimuli used in the EP paradigm were also used in the IAT. This post-hoc analysis should reveal whether the variable Stimulus IAT (-1 = stimuli not in the IAT, 1 = stimuli used in the IAT) further moderated the prime type by target type by *D* score interaction found for participants without LP. If an interaction of prime type, target type and *D* was only existent for trials in the EP paradigm for which the stimuli were also used in the IAT, then this would undermine the present interpretation of automatic activation of evaluations of ambiguity and clearness. As within one trial of the EP paradigm, the prime or the target or both could be stimuli used in the IAT, we only used trials of the EP paradigm for which both prime and target were stimuli used in the IAT or both stimuli were not used in the IAT. Mixed trials for which only the prime or the target stimulus was used in the IAT were not used for this analysis. Obviously, this analysis is based on the same number of participants but used less trials than the original analysis.

Table 23 shows the fixed effects of the mixed effects model with the additional effect coded variable Stimulus IAT. As this analysis was only done to test whether this variable further moderated the prime type × target type × *D* score interaction found for participants without LP, we only refer to this effect and the highest order effect. As can be seen in Table 23 the interaction of target type, prime type, *D*, LP and Stimulus IAT was not significant. This means that the target type × prime type × *D* × Stim IAT did not differ between participants with and without LP. More importantly, the interaction of target type, prime type, *D* and Stimulus IAT for participants without LP (dummy coded) was not significant. This means that regardless of whether the stimuli used in the EP paradigm were also used in the IAT or not, the target type by prime type by *D* interaction was significant.

Table 23

Fixed effects of mixed-effect model with mean centered D score including factor coding whether the stimuli used in the EP paradigm were also used in the IAT or not (study 3).

	b	*SE b*	*df*	*t*	*p*
Intercept	6.4124	0.0160	264	401.08	< .001
Target type N-1	0.0104	0.0009	48845	11.37	< .001
Target type	0.0252	0.0073	19	3.46	< .01
Prime valence	-0.0005	0.0010	35	-0.49	0.63
Prime type	0.0002	0.0023	467	0.09	0.92
D	0.0023	0.0285	438	0.08	0.94
Stimulus IAT (Stim IAT)	-0.0152	0.0073	19	-2.09	0.05
Learning Phase (LP)	-0.0071	0.0198	438	-0.36	0.72
Target type N-1 × Target type	-0.0173	0.0009	48855	-18.83	< .001
Target type × Prime valence	0.0042	0.0009	48833	4.62	< .001
Target type × Prime type	-0.0052	0.0022	48828	-2.38	0.02
Prime type × *D*	-0.0051	0.0043	48831	-1.17	0.24
Target type × *D*	0.0022	0.0043	48829	0.52	0.61
Prime type × Stim IAT	-0.0022	0.0022	464	-0.99	0.32
Target type × Stim IAT	0.0193	0.0073	19	2.65	0.02
D × Stim IAT	0.0040	0.0043	48825	0.92	0.36
Prime type × LP	-0.0002	0.0030	48840	-0.06	0.95

Target type × LP	-0.0038	0.0030	48826	-1.28	0.20
D × LP	-0.0472	0.0396	438	-1.19	0.23
Stim IAT × LP	-0.0017	0.0030	48836	-0.59	0.56
Target type × Prime type × *D*	-0.0124	0.0043	48828	-2.86	< .01
Target type × Prime type × Stim IAT	-0.0028	0.0022	48824	-1.28	0.20
Prime type × *D* × Stim IAT	-0.0056	0.0043	48824	-1.29	0.20
Target type × *D* × Stim IAT	0.0019	0.0043	48823	0.45	0.66
Target type × Prime type × LP	0.0078	0.0030	48827	2.62	0.01
Prime type × *D* × LP	-0.0002	0.0060	48844	-0.04	0.97
Target type × *D* × LP	-0.0099	0.0060	48828	-1.66	0.10
Prime type × Stim IAT × LP	0.0035	0.0030	48836	1.18	0.24
Target type × Stim IAT × LP	-0.0017	0.0030	48820	-0.58	0.56
D × Stim IAT × LP	-0.0015	0.0060	48840	-0.25	0.80
Target type × Prime type × *D* × Stim IAT	-0.0030	0.0043	48824	-0.70	0.48
Target type × Prime type × *D* × LP	0.0199	0.0060	48828	3.32	< .001
Target type × Prime type × Stim IAT × LP	0.0044	0.0030	48821	1.49	0.14
Prime type × *D* × Stim IAT × LP	0.0049	0.0060	48839	0.83	0.41
Target type × *D* × Stim IAT × LP	-0.0047	0.0060	48822	-0.79	0.43
Target type × Prime type × *D* × Stim IAT × LP	0.0030	0.0060	48823	0.51	0.61

Note. Effect coding was used for the dichotomous variables: Target type and Target type N-1 (-1 = positive, 1 = negative), Prime type (-1 = unambiguous, 1 = ambiguous), Stimulus IAT (-1 = stimuli not used in the IAT, 1 = stimuli used in the IAT). Prime valence was centered by the mean. Learning phase (0 = without, 1 = with). The *D* scores were separately centered by the sample mean for participants with and without LP. This analysis was based on 49,924 observations.

3.5.4 Discussion

In this study we tried to strengthen the evidence for interindividual differences in the automatic activation of evaluations of lexically ambiguous and unambiguous words via a conceptual replication of study 2. Furthermore, by adding a condition without learning which word is ambiguous or not prior to the EP paradigm we tried to expand the external validity in a sense that this condition allows to draw conclusions about whether the mental state of having two meanings vs. one meaning in mind (which is necessary for subjective ambiguity) triggers an (interindividually different) automatic activation of evaluation. Indeed, the prime type × target type × IAT *D* value interaction in the condition without learning phase provides evidence for this. However, the replication of the effect for participants with a learning phase was not successful. In the following, we will first focus on this failed replication with reference to procedural differences between the two studies. Then we will have a closer look at the significant three-way interaction in the condition without learning phase.

We changed the stimuli, number of trials, and the use of the built-in penalty in the IAT. All that could potentially have an impact and could contribute to the failed replication of the effect in the condition with a learning phase. However, we changed the learning phase itself. In study 2, the learning of (un)ambiguousness of the words was achieved by presenting each word (highlighted by underlining) in two sentences which should have made the different meanings salient for the ambiguous words. The task for the participants in that study was to decide whether the underlined word has one or two meanings. In the current study, participants were supposed to categorize the words according to their number of meanings (1 or 2) without contextual information. It could be that this procedural change increased the attention for the (un)ambiguousness of the primes which were presented in an isolated fashion instead of integrated in sentences like in the previous study. If the processing of the ambiguousness of the primes was indeed boosted in the current study, there are several possibilities of explaining the lack of prime type × Type × *D* interaction.

One possible post-hoc explanation comes from the segmentation hypothesis (Fiedler et al., 2011) which postulates that the prime loses the impact on the target evaluation if the prime is functionally separated from the target (e.g. due to an explicit response to the prime). In one experiment, Fiedler et al. (2011; Experiment 1c) could show that even by responding to a valence-irrelevant dimension of the prime, the standard congruity effect was eliminated. Deciding which in isolation presented word is ambiguous or not prior to the EP paradigm could result in a processing style in which participants classified the primes in terms of ambiguity as well. If that is the case, the classification of the ambiguousness of the prime could establish a segment boundary which prevents the primes from affecting the target evaluation.

Another explanation could be based on Alexopoulos et al. (2017), who provided evidence that the processing style (dissimilarity orientated or local focus vs. similarity or global focus) activated in a task prior to the EP paradigm can carry over into the EP paradigm and prevent the prime from exerting influence on the target evaluation. The authors interpreted their results in terms of the segmentation hypothesis. The learning paradigm in the present study (using isolated presented ambiguous or unambiguous words instead of words in sentences) potentially triggers a more a local focus within the EP paradigm than the sentences paradigm used in the previous study. There, the task was to decide if the same underlined word in two sentences displayed together on the screen had one or two meanings. One could argue that comparing the meaning(s) of a word which can only be done by taken the context into account triggers a more global focus than simply deciding about the number of meanings of one word. Hence, a potentially more local orientated processing style could lead to a segmentation of prime and target which (according to the segmentation hypothesis) in turn leads to a lack of prime by target interaction and no moderation of this interaction by the attitude towards ambiguity.

The feature-specific attention allocation model (Spruyt, De Houwer, Everaert, & Hermans, 2012; Spruyt, De Houwer, Hermans, & Eelen, 2007; Spruyt, Houwer, & Hermans, 2009) could also explain the lack of an effect, assuming that due to the learning paradigm more attention was aligned to the (un)ambiguousness of the primes. They found that semantic priming (including affective priming) emerged only if attention was aligned to that specific semantic stimulus feature. Although, according to the model, the evaluative processing of the target in our paradigm should result in an attentional focus on the evaluative dimension, the learning phase in this study prior to the EP paradigm put an attentional focus on the ambiguousness of the primes. The attentional focus was potentially stronger in the present study compared with the previous one, as the search for meanings in the present learning phase had to be done without additional information and was therefore probably more demanding. The potentially higher attention to the dimension of ambiguity (if it is transferred to the EP paradigm) could result in less attention to the task-irrelevant evaluative dimension of the prime.

The mentioned explanations for the discrepancy between this and the previous study could potentially be transferred to explain the marginally significant four-way interaction of prime type, target type, IAT *D* value, and learning phase, too. In the condition without a learning phase, participants probably did not focus on the ambiguousness of the primes (or at least less than participants with a learning phase), making it more likely that the valence due to the (un)ambiguousness could have an effect on the response to the target.

We will now take a closer look at the prime type × target type × IAT *D* value interaction for participants without a learning phase. In contrast to study 2, the mean *D* score with 0.31 was not equal to zero. If the IAT scores are seen as purely reflecting associative processes (although there is strong contrary evidence; Calanchini et al., 2014; Meissner & Rothermund, 2013), the IAT scores could be interpreted as a relative measure reflecting the liking of two concepts. In this view, a zero IAT score would indicate no preference of ambiguity over clearness or vice versa. According to this assumption, when analyzing the moderated prime type × target type interaction we could look at 1 *SD* below and above zero. Thus, we could see if the prime type × target type interaction reflects a relative preference of ambiguity over clearness for an uncentered *D* score of zero - 1 SD on the one hand and a relative aversion of ambiguity compared to clearness for an uncentered *D* score of zero + 1 SD on the other hand. This analysis revealed a significant prime type × target type interaction (in the predicted direction) only for participants for which the IAT indicated a relative liking of ambiguity over clearness. It is important to note that in study 2 the interaction of prime type by target type was significant only for participants for which the IAT indicated a relative liking of clearness over ambiguity (see right plot of Figure 11). If the IAT scores are considered as being driven by associative and non-associative processes (e.g. recoding by salience, ability to overcome the automatic association, task switching ability; Back et al., 2005; Calanchini et al., 2014; Rothermund & Wentura, 2004) a zero IAT *D* score cannot be taken as reflecting equal liking/disliking of the two target concepts. One possibility to deal with this is to use the information of the relative position of a given *D* compared to the average *D* score of the investigated sample (see Figure 13). Using the mean centered *D* scores revealed a marginally significant prime type × target type interaction for the mean centered *D* – 1 SD. The prime type × target type interaction for mean centered *D* + 1 SD was, at least descriptively, in the predicted direction. Indeed, the general pattern for the predicted latencies as a function of prime type and target type for low and high values of the *D* score using the mean centered IAT *D* scores (vs. the uncentered IAT *D* scores) was closer to the predicted pattern.

If participants were aware of the distinctive feature of the primes, those who evaluate ambiguity and situations in which they are confronted with ambiguity negatively could get in a negative mood as they are confronted with ambiguity within the EP paradigm. That could be problematic as it is known that negative mood inhibits EP effects (Storbeck & Clore, 2008b). While the interpretation of the authors was based on the mood modulated accessibility of semantic concepts, another possible explanation is a differential mood related segmentation of prime and target episodes (Alexopoulos et al., 2017). If indeed negative mood constrains the accessibility of semantic concepts, then this could prevent the mental state of having two meanings in mind and thus potentially eliminate possible EP effects due to the differential valence of having one vs. two meanings in mind. If the accessibility of the primes is highlighted via a learning paradigm, it could be assumed that the pre-activation of concepts reduces the impact of impaired accessibility of semantic concepts for participants in a negative mood. Indeed, our data were in line with this possibility. In this study, participants who had no learning phase prior to the EP paradigm and for whom the IAT indicated a relatively more negative attitude towards ambiguity than clearness (and hence could be in a negative mood due to the processing of ambiguity) showed no prime type by target type interaction. In contrast, in study 2 in which the accessibility of the meanings of the primes was highlighted via a learning paradigm prior to the EP paradigm, participants for whom the IAT indicated a relatively more negative attitude towards ambiguity than clearness showed a prime type by target type interaction.

If we consider the three-way interaction as reflecting a true effect, the question arises whether it is justified to draw conclusions about the interindividual automatic activation of evaluations of the (un)ambiguous primes. We think that this is only the case when one assumes that the IAT as well as the EP paradigm are at least partially influenced by the construct of interest: attitude towards ambiguity. Assuming that an attitude towards ambiguity does not exist, there would be no automatic activation of evaluations/valence, and the IAT effect would only be influenced by construct-unrelated variance. To the extent that the EP paradigm was also affected by such construct-unrelated variance, the three-way interaction would not be indicative of an automatic activation of evaluation of the phenomenon ambiguity or clearness. In a similar vein, Calanchini et al. (2014), in referring to their results (namely that the D and the OB parameter of the Quad model - see section 2.3.2 - seem to be relatively general non-attitudinal processes), mentioned that if the same general non-attitudinal processes influencing the IAT also influences other behavior (in our case, responses in the EP paradigm), a correlation without contribution of the attitude can occur. Obviously, the (assumedly stable) ability to overcome the bias (OB) of the automatically activated evaluation/association influences both the EP paradigm and the IAT in a similar way. There is evidence that higher OB leads to smaller IAT effects (Conrey et al., 2005) as well as to smaller EP effects (Allen, Sherman, & Klauer, 2010). In principle, this could partly explain a simple IAT-EP correlation (the same could be applied to the *D* parameter which reflects the ability to detect the correct response) but not our predicted, more complex three-way interaction, which makes not only predictions about the size of the EP effect depending on the IAT *D* score but also about the directions of the prime type by target type interaction. The latter seems to be at least descriptively confirmed regarding the analysis with a mean centered IAT *D* score.

However, there could be another source of unintended construct-unrelated variance affecting both the IAT and the EP paradigm in such a way that the three-way interaction emerges without contribution of automatic activation of evaluation of ambiguity but rather as a result of the automatic activation of evaluation of the individual words, which have idiosyncratic positive or negative valence although they were selected to be neutral. Assuming that the list of ambiguous and clear words randomly differs in valence (not systematically affected by the distinction between ambiguity and clearness), then the IAT score probably will be affected by this difference. Especially when the same words or a subset of words are used in the EP paradigm, a moderation of the prime type by target type interaction by the IAT score could not be interpreted in terms of automatic activation of evaluations of the phenomenon ambiguity and clearness but due to this random variation in valence over the lists of ambiguous and clear words. Obviously, this possible explanation seems unlikely in light of the view that the IAT effect in general is predominantly driven by influences on the concept level and less driven by the valence of the individual stimuli (see 2.3.2). However, De Houwer (2001; see footnote 4) mentioned that the valence of the individual stimuli influenced the IAT effect when the target concepts did not differ in valence. As this could be true for our target concepts as well, we conducted a post hoc analysis in order to exclude such an alternative explanation for our results. That is, we included an additional factor in the model distinguishing between words used in the EP paradigm as well as the IAT and word not used in both tasks. This analysis (see Table 23) revealed that the three-way interaction was not further moderated by this factor. Although the number of trials was reduced, the prime type × target type × *D* score interaction remained significant. This post-hoc analysis strengthens the view that there are interindividual differences in the automatic activation of evaluations of having multiple meanings in mind.

As a side note, it is worth looking at the target type × prime valence interaction (included for error binding reasons), which represents the usual EP effect (faster latencies if prime and target share the same valence). This highly significant interaction can be taken as an indicator of adequate data quality. Another side note needs to be mentioned regarding the strong interaction of target type with the previous target type. This is probably driven by the fact that we did not disable the possibility that two targets words could be identical in two consecutive trials (5 % by chance), probably resulting in faster word recognition. In subsequent experiments we excluded identical word repetition in order to reduce this methodological source of unintended variance.

Notwithstanding the discrepancies to study 2, the current study provided some evidence for interindividual differences in the automatic activation of evaluations of the mental state of having two meanings vs. one meaning in mind, which we consider to be necessary for speaking about subjective ambiguity. In order to obtain more certainty about the effect, we ran a replication of the condition without a learning phase.

3.6 Study 4

This pre-registered study (see Appendix F) was a partial replication of study 3. We attempted to replicate the evidence in study 3 of automatic activation of valence of lexically ambiguous and unambiguous words assessed with the EP paradigm (in line with the measured associations of the concepts of ambiguity and clearness with valence assessed with the IAT) for lexically (un)ambiguous words for which there was no learning of the meanings of the words prior to the EP paradigm. Therefore, we ran a direct replication of the condition without learning phase prior to the EP paradigm of study 3.

3.6.1 Hypothesis

As in study 3 we expected that the prime type × target type interaction is moderated by the IAT *D* value (see Figure 8).

3.6.2 Method

Participants. Data were collected for one week from May 5 to May 12, 2020. All participants were recruited via the mailing list of the University of Tübingen. We only included data from participants who finished the study completely. This resulted in valid data sets from 408 participants (322 females, 84 males, 2 diverse) with a mean age of 25.92 years (*SD* = 9.71 years). Unlike study 3, at the beginning of this study participants had to actively confirm that they speak German on native speaker level and that they did not use a smartphone for participating in the study in order to get access to the study. Therefore, this time we did not have to exclude participants with potentially missing German language skills or those who used a smartphone. All participants participated in exchange for course credit or could take part in a lottery for one of ten 20 Euro vouchers for a chocolate shop located in Tübingen.

Design. A 2 (prime type: ambiguous vs. unambiguous) × 2 (target type: positive vs. negative) repeated measure design was used.

Materials. Stimuli were identical to those used in study 3.

Procedure. The online study used the JavaScript library jsPsych (de Leeuw, 2015). It ran in full-screen mode with black font on white background on the PC or laptop of the participants. After the demographic questions the EP paradigm was done with the same instructions and number of blocks and trials per block as in study 3. Timing parameters (SOA: 300 with 200 ms prime presentation) were identical except that the ITI was 800 ms instead of 810 ms and the 10 ms blank screen between fixation cross and the prime was removed. Another difference between the previous studies and this one was that a repetition of the same prime or target word in two consecutive trials was disabled. The IAT was the same as in study 3. At the end, participants could provide information (not saved together with previous data) for the lottery or for course credit.

3.6.3 Results

Data trimming and the applied multilevel model (see Appendix F) was the same as in study 2 and, except for the factor learning phase also the same as in study 3. We excluded all trials in the EP paradigm with reaction times below 300 ms (0.47 %) and above 3000 ms (0.17 %) and all false classifications (4.70 %) from the analysis. In total we excluded of 5.08 % of EP paradigm trials. This resulted in 92,946 remaining valid trials for the EP paradigm.

Again, like in study 3 the IAT had a built-in penalty. The improved algorithm was used to calculate the *D* score. The time to respond to the correct response key after making a mistake took *M* = 832 ms (median = 568 ms). A positive *D* score indicated that ambiguity was more strongly associated with negativity and clearness more with positivity than ambiguity with positivity and clearness with negativity. The IAT data of 9 participants (those who responded in less than 300 ms in more than 10 percent of trials) were excluded. Therefore, a *D* score was calculated for 399 participants. The mean *D* score was 0.38 (*SD* = 0.36). The odd even reliability for the *D* score was .66. The latencies in ms (derived from the application of data trimming according to the improved algorithm for the *D* score) for the congruent block (*M* = 1023, *SD* = 279.70) were faster than the latencies for the incongruent block (*M* = 1259, *SD* = 358.90), $t(398) = 17.32$, $p = <.001$.

The specified mixed-effect model was based on 90,228 observations from 399 participants (319 females, 78 males, 2 diverse). They had a mean age of 25.95 years (*SD* = 9.74 years). Random effect estimation for variance (*SD*) for the intercepts of participants was 0.0201 (0.1437), for the target stimuli 0.0017 (0.0412) and for the prime stimuli 0.0000 (0.0035). Fixed effects of the model with an uncentered *D* score can be seen in Table 24.

Table 24

Fixed effects of mixed-effect model with uncentered D score (Study 4)

	b	*SE b*	*df*	*t*	*p*
Intercept	6.4283	0.0140	87	460.05	< .001
Target type N-1	0.0054	0.0007	89799	8.22	< .001
Target type	0.0215	0.0093	18	2.32	.03
Prime valence	-0.0012	0.0008	37	-1.40	.17
Prime type	-0.0003	0.0011	100	-0.26	.79
D	-0.0214	0.0201	397	-1.06	.29
Target type N-1 × Target type	0.0016	0.0007	89796	2.35	.02
Target type × Prime valence	0.0025	0.0006	89767	3.86	< .001
Target type × Prime type	-0.0023	0.0010	89767	-2.37	.02
Prime type × *D*	0.0000	0.0018	89766	0.02	.99
Target type × *D*	0.0012	0.0018	89767	0.65	.52
Target type × Prime type × *D*	-0.0020	0.0018	89767	-1.10	.27

Note. Effect coding was used for the dichotomous variables: Target type and Target type N-1 (-1 = positive, 1 = negative), Prime type (-1 = unambiguous, 1 = ambiguous). Prime valence was centered by the mean. The *D* score was uncentered (the mean *D* score was 0.38).

Like the previous studies, the target type N-1 effect showed longer latencies when the previous trial was negative (vs. positive). In line with the literature, the target type effect revealed longer latencies in the case of negative (vs. positive) targets (Dijksterhuis & Aarts, 2003; Unkelbach et al., 2008). The target type N-1 × target type interaction was significant (but much weaker than in the previous studies in which the random repetition of the same prime or target word in two consecutive trials was not disabled) and showed, like in previous studies, shorter latencies when the previous and the current target type had the same valence. Again, as in study 2, we found a significant target type × prime type interaction in the analysis without centering the *D* score. That means that for a *D* score of zero (for which the IAT *D* score indicated that there is no difference between the strength of associations of the concepts of ambiguity and clearness with positive and negative valence) the influence of prime type differed between the positive and negative targets. The influence of prime type for negative targets (b = -0.0026, $SE\ b$ = 0.0015, $t(302) = -1.75$, $p = .08$) was marginal significant. That means that descriptively shorter latencies for negative targets were observed when on that trial the primes where ambiguous (vs. clear). The influence of prime type for positive targets was not significant (b = 0.0020, $SE\ b$ = 0.0015, $t(309) = 1.35$, $p = .18$). The replicated result of target type × prime type interaction in the analysis without centering the *D* score and the expected but not significant three-way interaction will be discussed later. Although the three-way interaction is not significant, predicted values for -1 SD (+1 *SD)* below (above) the uncentered *D* score of zero can be found in Figure 14.

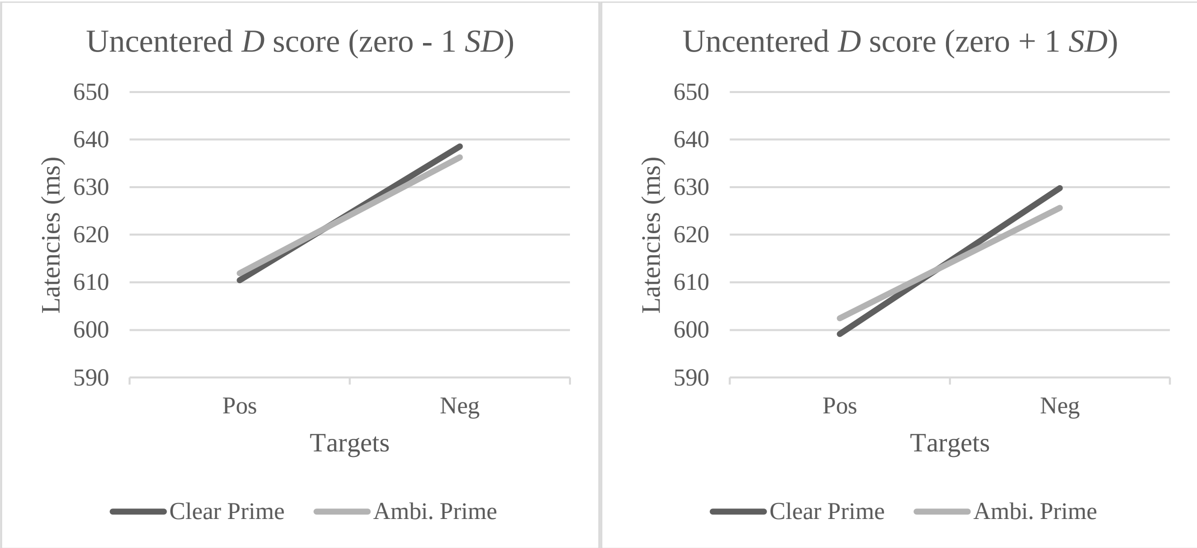


Figure 14. Left (Right) plot shows the predicted values for -1 SD (+1 *SD)* below (above) the uncentered *D* score, that is for participants for which the IAT indicated that they like ambiguity more than clearness (clearness more than ambiguity).

Nonetheless, further analyses showed that the prime type × target type interaction was not significant for the *D* score of zero – 1 *SD* (left plot of Figure 14), *b* = -0.0015 (SE *b* = .0015), *t*(89767) = -1.03, p = .31, but highly significant for the *D* score of zero + 1 *SD* (right plot of Figure 14) with *b* = -0.0030 (SE *b* = .0007), *t*(89767) = -4.53, *p* < 001. It showed longer latencies given ambiguous (vs. clear primes) when the target was positive (*b* = 0.0027 (SE *b* = .0011), *t*(94) = 2.49, *p* = .01) but shorter latencies given ambiguous (vs. clear primes) when the target was negative (*b* = -0.0033 (SE *b* = .0011), *t*(92) = -3.02, *p* < .01). The prime type × target type interaction for the *D* score of zero + 1 *SD* (right plot of Figure 14) was in line with our hypothesis: A positive *D* score indicated a relative disliking of ambiguity compared with clearness. Therefore, ambiguous vs. clear primes probably triggered stronger negative evaluations which were evaluatively congruent (incongruent) with negative (positive) targets, resulting in faster (slower) latencies.

As already discussed in the previous studies, a zero IAT *D* score (as influenced by associative and non-associative processes) does not necessarily reflect equal liking/disliking of ambiguity and clearness. Therefore, an alternative option is to use the information of the relative position of a given *D* within the distribution of *D* scores in the sample. This can be done by centering the *D* scores by the sample mean. As the right plot of Figure 14 shows, the predicted latencies as a function of prime type and target type for 1 *SD* (.36) above the uncentered *D* score of zero were almost identical to the predicted values using the mean (*M* = 0.38) centered *D* scores. Obviously the prime type × target type interaction for the mean centered *D* score was almost identical with $b = -0.0030$ (SE $b = .0007$), $t(89767) = -4.56$, $p < .001$ to the interaction of prime type by target type for the uncentered *D* score zero + 1 *SD* (left plot of Figure 14). Likewise, the simple effects of prime type for positive and negative targets are almost identical to the simple effects reported above for the interaction of prime type by target type for the uncentered *D* score zero + 1 *SD*. Therefore, we did not report these effects. As the main effects and interactions (except for the prime type by target type interaction which was stronger for the mean centered *D* score analysis compared to the uncentered analysis) was almost identical to the reported effects in Table 24, we did not comment on these effects. The implications of a highly significant prime type × target type interaction for the mean centered *D* score will be discussed later.

3.6.4 Discussion

The aim of this replication study was to add evidence for interindividual differences in the automatic activation of evaluations of ambiguity/clearness via replicating the prime type by target type by IAT *D* score interaction of Study 3. This was not successful. However, we found a strong prime type × target type interaction for an IAT *D* score of zero derived from the uncentered as well as from a mean centered version of this variable. This interaction indicated that ambiguous primes automatically activated more negative (or less positive) valence than unambiguous primes.

As already discussed in study 2, a significant prime type × target type interaction for an IAT *D* score of zero derived from the uncentered version of this variable is conter-intuitive and needs to be discussed. We already mentioned that there is strong evidence that the performance in the IAT (as well as the EP paradigm) is driven by associative and non-associative processes. Therefore, a zero IAT *D* score does not necessarily reflect equal liking/disliking of the two target concepts. This could potentially explain this inconsistency. We consider this repeated finding with different and more stimuli as additional evidence for the assumption that the neutral points for which each measure indicates no relative preference of one concept over the other were not identical.

The (even stronger) highly significant prime type × target type interaction for an zero IAT *D* score derived from the mean centered version of this variable (for visualization see right plot of Figure 14, as the predicted values for an uncentered *D* score + 1 *SD* were almost identical to the predicted values for the mean centered *D* scores, as the *SD* is almost equal to the *M*), provided evidence that ambiguous primes automatically activated more negative or less positive valence than unambiguous primes for people with an average attitude towards ambiguity/clearness (measured by the IAT). This pattern is in line with the average *D* score of 0.38 which indicates a relative preference for clearness over ambiguity.

It could be speculated whether the period of data collection of this study had an impact on our results. The study was conducted during the first wave of the COVID-19 crisis which led to the experience of high levels of uncertainty in the society (Rettie & Daniels, 2020). This exceptional situation could have led to a higher desire for clear, structured, non-contradictory and unambiguous information and to higher levels of aversion to uncertainty and ambiguity. This potential COVID-19 crisis-induced attitudinal change for ambiguity and clearness could have led to the strong prime type × target type interaction. On the other hand, comparing the size of the *D* score with the *D* scores of the three previous studies does not provide evidence for a strong attitudinal change.

To sum up, there is at least evidence for the automatic activation of evaluations of ambiguity/clearness (as indicated by the prime type by target type interaction) but with this study we could not provide further evidence that these evaluations of ambiguity and clearness differ between participants depending on the attitude towards these concepts (as the prime type × target type × IAT *D* score interaction was not significant).

As we did not know which amount of variance of the ambiguity-IAT *D* scores was determined by associative processes and which amount was driven by non-associative processes, we could not rule out insufficient construct related variance in this sample as an explanation for the non-significant three-way interaction. Therefore, we decided to counteract this possibility by experimentally manipulating the attitude towards ambiguity and clearness prior to the EP paradigm in the following two studies.

3.7 Study 5

With this pre-registered study (see Appendix G1) we further investigated the assumed automatic activation of valence of lexically ambiguous and unambiguous words depending on the attitude towards the concepts of ambiguity and clearness. Again, automatic activation of valence was assessed via the EP paradigm. Instead of measuring the attitude towards these concepts (via the IAT) and using this information as a moderator of the prime type × target type interaction, we tried to manipulate the attitude towards these concepts prior to the EP paradigm. Ebert, Steffens, Stülpnagel, and Jelenec (2009) provided evidence that the IAT changes attitude via associative learning when using only one half of the IAT. Associative learning should occur as conducting an IAT includes spatio-temporal continuity of targets together with attributes at the level of the mental representations. Using the first three blocks of the IAT we tried to induce associations of the concepts of ambiguity and clearness with positive or negative valence, respectively. In order to investigate if the potential learning of associations is bound to specific stimuli (those used in the IAT) or generalized to concepts and therefore to new ambiguous and unambiguous words, too, we used two different stimulus sets for the manipulation phase.

3.7.1 Hypothesis

(1) We expected that the prime type × target type interaction is moderated by the manipulated attitude towards the concepts of ambiguity and clearness (see Figure 15). In the ambiguity = negative and clearness = positive condition, we expected that if an ambiguous prime precedes a negative target, the reaction time should be faster than if an unambiguous prime precedes a negative target. If an ambiguous prime precedes a positive target, the reaction time should be slower compared to an unambiguous prime preceding a positive target. In the ambiguity = positive and clearness = negative condition, the prime type × target type interaction should indicate the reversed pattern (cf. Appendix G1).

(2) Furthermore, we expected that the three-way interaction postulated in (1) is stronger for stimuli used in the manipulation phase than for stimuli not used in the manipulation phase.

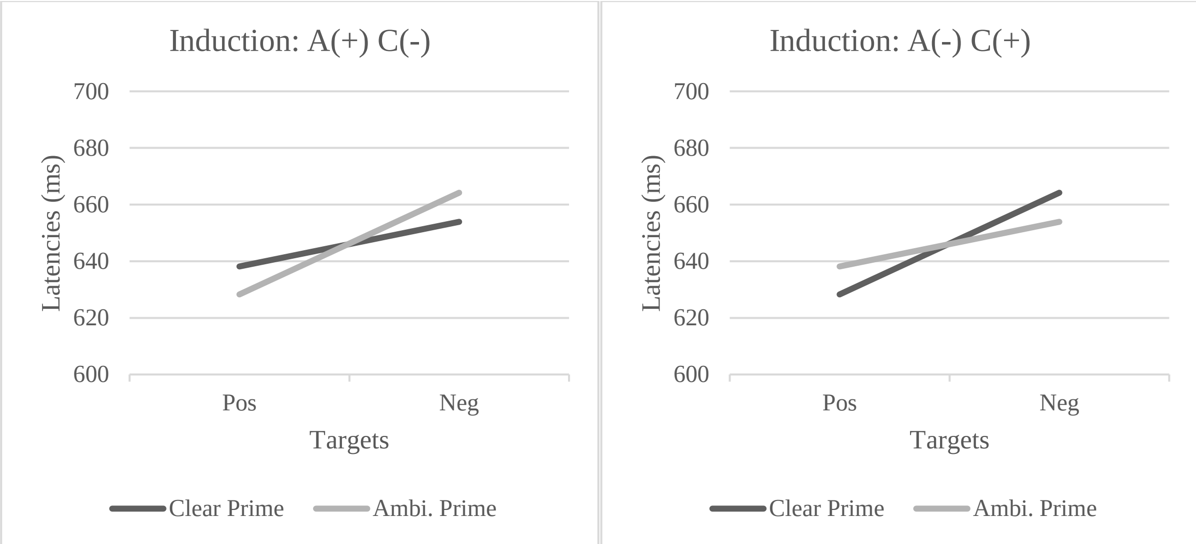


Figure 15. Predicted pattern of interaction of Prime type and target type by Induction. A(+) C(-) means that in the induction phase ambiguity was paired with positive and clearness was paired with negative and A(-) C(+) means ambiguity was paired with negative and clearness was paired with positive.

3.7.2 Method

Participants. Data were collected for one week from February 18 to February 25, 2019. All participants were recruited via the mailing list of the University of Tübingen. We only included data from participants who finished the study completely. Only participants who actively confirm that they speak German on native speaker level and that they did not use a smartphone for participating the study could take part in the study. We collected full data sets from 293 participants (209 females, 77 males, 5 diverse and 2 with missing gender information) with mean age of 27.28 years (*SD* = 10.83 years). 141 (152) participants were randomly assigned to the attitude manipulation condition ambiguous = positive + clearness = negative (ambiguous = negative + clearness = positive). The stimulus set for the manipulation phase was counterbalanced. 162 (131) participants had stimulus set A (B). All participants participated in exchange for course credit or could take part in a lottery for one of ten 20 Euro vouchers for a chocolate shop located in Tübingen.

Design. We used a 2 (induced association: ambiguity = positive + clearness = negative vs. ambiguity = negative + clearness = positive) × 2 (prime type: ambiguous vs. unambiguous) × 2 (target type: positive vs. negative) × 2 (stimulus novelty: same as in induction phase vs. different from induction phase) design with repeated measures on the last three factors. Additionally, across participants we counterbalanced the specific stimulus set used for the manipulation phase.

Materials. The NAUG database (see Appendix A3 – A5) served as basis for selection of stimulus words. The matching was done with the program Match (van Casteren & Davis, 2007). Prior to the matching process we preselected ambiguous and unambiguous words. This preselection (exclusion of unbalanced ambiguous words, words with more than two dominant meanings and ambivalent ambiguous words) was identical to the one in study 3. From the unambiguous words all strongly positive (mean rating above 7) and all strongly negative (mean rating below 3) words were used for the matching of positive and negative words by all other dimensions. The remaining unambiguous stimuli were used for matching with the 61 preselected ambiguous stimuli. In a first step an item by item matching was done with the unambiguous items. Then the 20 most neutral ambiguous stimuli were selected together with the corresponding matched unambiguous stimuli and split in two equal sized item sets with similar valence distribution and mean valence. Because of unacceptable matching in the other dimensions these item sets had to be reduced until an acceptable match of all dimensions (word frequency, number of letters, valence, arousal and abstractness) was achieved. As in study 3, for the matching of valence, arousal and abstractness the means of the mean ratings of both meanings were used. This resulted in 2 stimuli sets (A and B) each containing 6 ambiguous and 6 unambiguous stimuli (see Appendix G2; including *M* and *SD*). The selected unambiguous positive and negative words were split in two item sets with similar distribution and mean valence. Then these two sets (A and B) were each reduced to the 6 positive and 6 negative words matched for all dimensions except for valence (see Appendix G3; including *M* and *SD*). Thus, stimulus sets A and B consisted of 6 ambiguous, 6 unambiguous, 6 positive, and 6 negative words each.

For stimuli set A (B) the 6 unambiguous and 6 ambiguous words showed a good matching in all dimensions: frequency, $t(10) = 0.12$, $p = .91$ ($t(10) = 0.27$, $p = .80$), number of letters, $t(10) = 0$, $p = 1$ ($t(10) = 1.24$, $p = .24$), valence of the word $t(10) = 0.30$, $p = .77$ ($t(10) = -1.96$, $p = .08$), mean valence of meanings $t(10) = 0.89$, $p = .39$ ($t(10) = 0.35$, $p = .73$), arousal of the word $t(10) = 3.40$, $p < .01$ ($t(10) = 1.13$, $p = .28$), mean arousal of meanings $t(10) = 2.11$, $p = .06$ ($t(10) = 0.81$, $p = .44$), abstractness of the word $t(10) = 0.66$, $p = .52$ ($t(10) = 0.27$, $p = .79$) and mean abstractness of meanings $t(10) = 0.81$, $p = .44$ ($t(10) = 0.28$, $p = .79$).

The matched 6 positive and 6 negative words of stimuli set A (B) did not differ in frequency, $t(10) = -0.81$, $p = .44$ ($t(10) = -0.86$, $p = .41$), number of letters, $t(10) = -0.43$, $p = .68$ ($t(10) = 0.40$, $p = .70$), arousal of the word $t(10) = -1.18$, $p = .27$ ($t(10) = -0.55$, $p = .59$), abstractness of the word $t(10) = -0.13$, $p = .90$ ($t(10) = -0.16$, $p = .88$) but as intended in valence $t(10) = 21.99$, $p < .001$ ($t(10) = 26.56$, $p < .001$).

Procedure. The online study used the JavaScript library jsPsych (de Leeuw, 2015). It ran in full-screen mode with black font on white background on the PC or laptop of the participants. After the demographic questions, the first three blocks of the IAT were conducted. We counterbalanced the specific set used for the IAT across participants. In the induced association condition ambiguity = positive + clearness = negative in the first block of the IAT the target concept clearness was displayed on the left and ambiguity on the right. The response keys were E (left) and I (right). In the condition ambiguity = negative + clearness = positive this pairing was switched. In the second block the attribute concepts negative (left) and positive (right) were learned. This localization was the same for both conditions. In the third block the target concepts and attribute concepts shared the previously learned response keys. Attribute stimuli (e.g., "SOMMER") and attribute concept labels (e. g., "POSITIVE") were presented in capital letters and target concepts and target stimuli with common lettering (e.g. "Zylinder", "Mehrdeutig). Block 1 and 2 had 24 trials. Block 3 had 48 trials. Inter trial interval was 250 ms. In order to continue after an incorrect response (which was displayed via a red X), the participants had to press the correct response key (built-in penalty).

After the manipulation-IAT the participants conducted the EP paradigm. Instructions were identical to previous studies and put a focus on fast responses. We used other stimuli in the practice block (20 trials with 300 ms feedback) than in the test blocks. The ambiguous and unambiguous primes as well as the positive and negative targets used in the practice block were matched for the same dimensions as the test stimuli. Stimulus sets A and B were mixed within each test block of the EP task. There were two test blocks (96 trials each). There was no feedback after the trials of the test blocks. Within each trial prime and target words belonged to the same stimulus set. Therefore, half of the stimuli used in the test blocks in the EP paradigm were known from the manipulation phase and half were new stimuli. Like in study 2 and 3, a trial consisted of a fixation cross (500 ms), a blank screen (10 ms), the prime (200 ms), a blank screen (100 ms) and the target which was displayed until a response was given. The ITI was 810 ms. A repetition of the same prime or target word in two consecutive trials was disabled. Between blocks participants could take a break. Response keys were E for negative words and I for positive words.

After the EP paradigm the manipulation check IAT was conducted. Participants always received the same stimuli as in the manipulation IAT. We used a seven block-IAT. Block 1, 2, 3 and 6 had 24 trials, the other blocks had 48 trials each. Initial target concept localization (left or right) was counterbalanced across participants. This established the sequence of IAT tasks: ambiguity = negative + clearness = positive (blocks 3+4) followed by ambiguity = positive + clearness = negative (6+7) or the reversed pattern. The attribute concept localization [negative (left), positive (right)] was fixed across participants. In other respects, this IAT was identical to the manipulation IAT at the beginning. At the end, participants could provide information (not saved together with previous data) for the lottery or for course credit.

3.7.3 Results

Manipulation-Check. As a manipulation check, we regressed the *D* scores of the manipulation check IAT on the induced association × sequence of IAT tasks interaction and the main effects. The *D* score was coded (irrespectively of sequence of IAT tasks) in a way that positive values indicated stronger ambiguity = negative and clearness = positive associations than ambiguity = positive and clearness = negative associations. As in study 3 and 4, RT of false responses included a built-in penalty (forced correct key press after wrong classification). The time to respond to the correct response key after making a mistake took a median of 690 ms (mean 512 ms) which is compatible to the 600 ms penalty used in study 2 which had no built-in penalty. The manipulation check IAT data of 10 participants of (those who responded in less than 300 ms in more than 10 percent of trials) were excluded. Hence, a *D* score for 283 participants was calculated. The odd even reliability for the *D* score was .84. Table 25 shows the estimated effects of the regression analysis.

Table 25

Estimated effects of regression analysis of the manipulation-check (study 5).

	b	*SE b*	*t*	*p*
Intercept	0.13	0.02	5.71	< .001
Induction	0.24	0.02	10.85	< .001
Sequence of IAT tasks	0.13	0.02	5.60	< .001
Induction × Sequence of IAT tasks	0.01	0.02	0.60	.55

Note. Effect coding was used: Induction (-1 = ambiguity + positive and clearness + negative, 1 = ambiguity + negative and clearness + positive), Sequence of IAT tasks (-1 = ambiguity + positive and clearness + negative followed by ambiguity + negative and clearness + positive, 1 = ambiguity + negative and clearness + positive followed by ambiguity + positive and clearness + negative). The *D* score is uncentered (the mean *D* score was 0.13). Multiple R-squared was 0.35, Adjusted R-squared was 0.34.

Figure 16 visualize the influence of induction and sequence of IAT tasks on the *D* score of the manipulation check IAT. The main effect of induction was significant. The induction of ambiguity + negative and clearness + positive led to higher *D* scores (*M* = 0.37, *SD* = 0.39) in the manipulation check IAT compared to the induction of ambiguity + positive and clearness + negative (*M* = -0.11, *SD* = 0.40). Furthermore, the main effect of Sequence of IAT tasks was significant. The sequence of IAT tasks with ambiguity + negative and clearness + positive followed by ambiguity + positive and clearness + negative revealed higher *D* scores (*M* = 0.26, *SD* = 0.45) compared to the sequence of IAT tasks with ambiguity + positive and clearness + negative followed by ambiguity + negative and clearness + positive (*M* = 0.01, *SD* = 0.44). The interaction effect of induction by sequence of IAT tasks was not significant. This means that the influence of induction did not differ between the two sequences of IAT tasks.

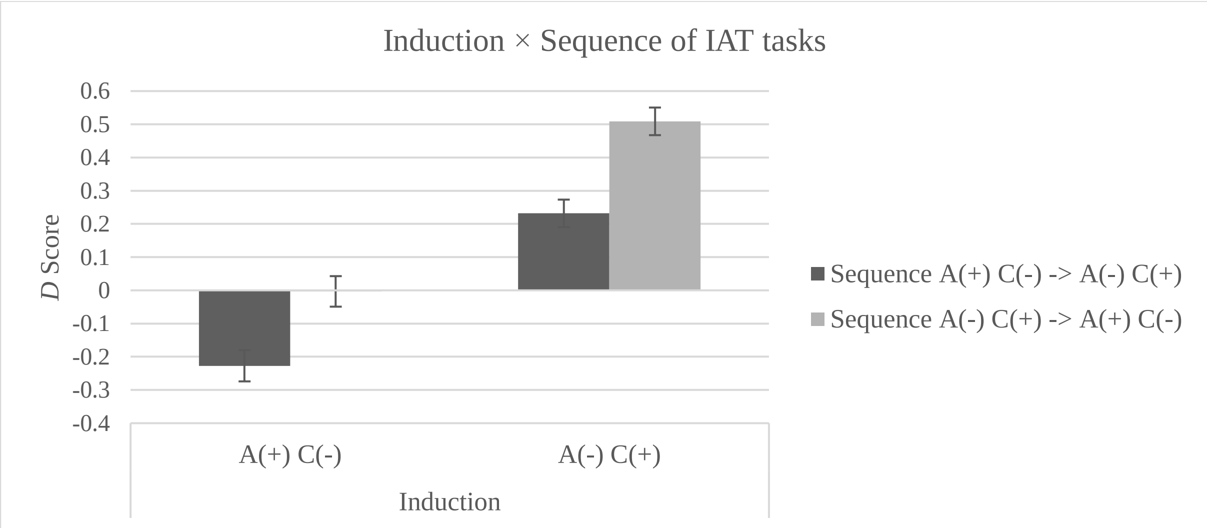


Figure 16. Influence of Induction and sequence of IAT tasks on the *D* score of the manipulation check IAT. A(+) C(-) means that ambiguity was paired with positive and clearness was paired with negative and A(-) C(+) means that ambiguity was paired with negative and clearness was paired with positive; Error bars represent standard errors of the mean.

Main analysis. We applied a multilevel model. The same random effects were estimated as in the previous studies. In the fixed effects model, we regressed the logarithmized reaction times in the EP paradigm on prime type × target type × induced association × stimulus novelty × stimulus set used in the induction phase and all subordinate four-way, three-way, two-way interactions and the main effects. Like in previous studies, we included a two-way interaction of the target type with the target type of the previous trial. The target type of the previous trial as a main effect was also included. As the stimulus set (A or B) used in the induction phase did not interact with any other variable, we excluded it from the model.

As in previous studies, we excluded all trials in the priming paradigms with reaction times below 300 ms (1.15 %) and above 3000 ms (0.26 %) and all false classifications (4.31 %)

from the analysis. In total we excluded 5.12 % trials of EP paradigm. This resulted in 53,377 remaining valid trials for the EP paradigm.

Random effect estimation for variance (*SD*) for the intercepts of participants was 0.017 (0.1309), for the target stimuli 0.0011 (0.0330) and for the prime stimuli 0.0000 (0.0004). Fixed effects of the model can be seen in Table 26.

Table 26

Fixed effects of mixed-effect model (Study 5)

	b	*SE b*	*df*	*t*	*p*
Intercept	6.4532	0.0103	101	628.24	< .001
Target type N-1	0.0074	0.0009	52493	8.20	< .001
Target type	0.0227	0.0068	22	3.34	< .01
Prime type	0.0014	0.0013	21	1.08	.29
Induction	0.0112	0.0077	290	1.45	.15
Stimulus novelty (SN)	-0.0062	0.0009	52491	-6.90	< .001
Target type N-1 × Target type	0.0007	0.0009	52496	0.72	.47
Target type × Prime type	-0.0010	0.0009	52477	-1.13	.26
Prime type × Induction	-0.0004	0.0009	52475	-0.50	.62
Target type × Induction	-0.0002	0.0009	52476	-0.24	.81
Prime type × SN	0.0008	0.0009	49409	0.84	.40
Target type × SN	0.0002	0.0009	52491	0.25	.80
Induction × Stimulus novelty	-0.0004	0.0009	52498	-0.46	.65
Target type × Prime type × Induction	0.0020	0.0009	52475	2.21	.03
Target type × Prime type × SN	-0.0003	0.0009	52475	-0.32	.75
Prime type × Induction × SN	0.0009	0.0009	40422	1.01	.31
Target type × Induction × SN	0.0006	0.0009	52497	0.67	.50
Target type × Prime type × Induction × SN	0.0006	0.0009	52475	0.65	.52

Note. Effect coding was used for the dichotomous variables: Target type and Target type N-1 (-1 = positive, 1 = negative), Prime type (-1 = unambiguous, 1 = ambiguous), Stimulus novelty (-1 = new stimuli, 1 = stimuli from induction phase), Induction (-1 = ambiguity was paired with positive and clearness was paired with negative, 1 = ambiguity was paired with negative and clearness was paired with positive). This analysis was based on 52826 observations from 293 participants.

As can be seen in Table 26 the four-way interaction of target type, prime type, Induction, and Stimulus novelty was not significant. This means that the target type × prime type × Induction is not further moderated by stimulus novelty. Stimulus novelty only had a significant main effect, which means that the responses for stimuli known from the induction phase were faster than responses for new stimuli. As can be seen in Figure 17, the target type × prime type × Induction interaction showed a pattern that was almost opposite to predictions.

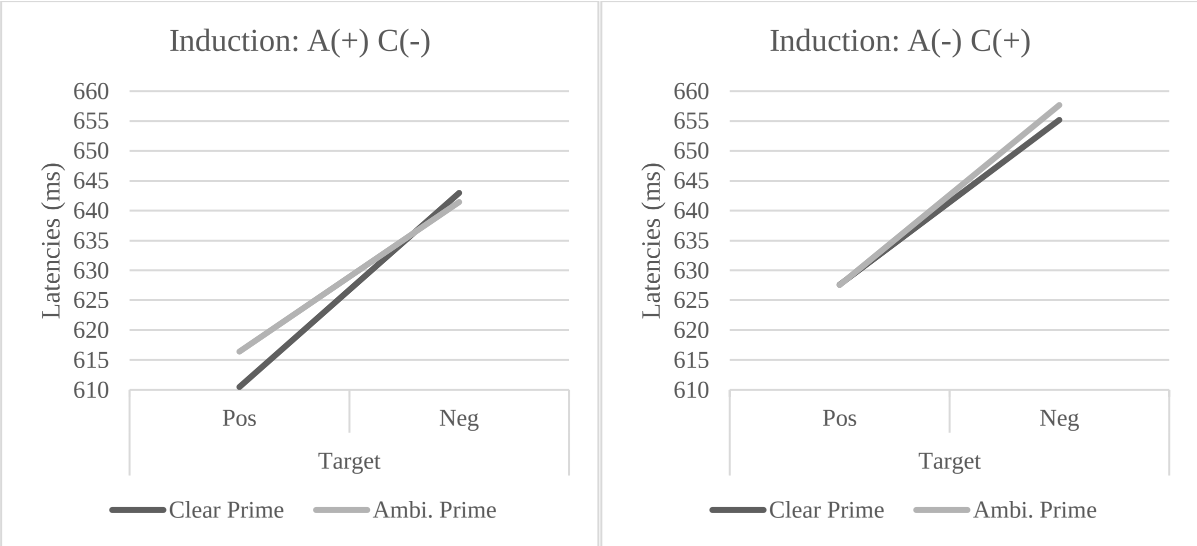


Figure 17. Interaction of prime type and target type by Induction. Left (Right) plot shows the predicted values for the induced association of ambiguity with positive and clearness with negative (ambiguity with negative and clearness with positive). As the relevant comparison were within-subject comparisons, we did not include standard errors of the mean.

We dummy coded the variable induction in order to further investigate the three-way interaction. The prime type × target type interaction for the induced association A(+) C(-) was significant (b = -0.0030, SE b = 0.0013, t = -2.31, p = .02) but not for the induced association A(-) C(+) (b = 0.0010, SE b = 0.0012, t = 0.78, p = .43). The simple main effect of prime type for positive targets for the induced association A(+) C(-) was significant with b = 0.0050, SE b = 0.0021, t = 2.35, p = .02. It showed that the latencies for positive targets were longer when ambiguous (vs. clear) primes preceded the targets. This was unexpected because the induced association of ambiguity (clearness) with positivity (negativity) should have led to a positive (negative) evaluation of ambiguous (clear) primes, thus facilitating (slowing) reactions to this target type. The simple main effect of prime type for negative targets for the induced association A(+) C(-) was not significant (b = -0.0011, SE b = 0.0021, t = -0.58, p = .56)

3.7.4 Discussion

This study investigated whether an automatic activation of valence of lexically ambiguous and unambiguous words occurred depending on the manipulated attitude towards the concepts of ambiguity and clearness. Furthermore, by counterbalancing the stimuli set used for the induction phase we investigated if the potential learning of associations is bound to specific stimuli or generalized to concepts. If generalization occurred new ambiguous and unambiguous words in the EP paradigm should automatically activate evaluations, too. The manipulation check suggested that the induction was successful. *D* scores for the induction of ambiguity + negative and clearness + positive (induction of ambiguity + positive and clearness + negative) were positive (negative). Moreover, the manipulation check IAT revealed that the induction of associations of ambiguity with negative and clearness with positive was more successful than the induction of associations of ambiguity with positive and clearness with negative. This can be interpreted in a way that the sample already had a stronger association of ambiguity with negative and clearness with positive. This is in line with the IAT results from studies 1, 3 and 4 which also indicated that clearness is preferred over ambiguity. The higher *D* scores of the sequence of IAT tasks with ambiguity + negative and clearness + positive followed by ambiguity + positive and clearness + negative compared to the reversed sequence could also be interpreted as evidence for generally stronger associations of ambiguity with negative valence and clearness with positive valence compared to the associations of ambiguity with positive valence and clearness with positive negative valence, because it is known that the IAT effects are stronger for IATs in which the congruent task precedes the incongruent task ((however, the adopted doubling of the practice trials in block 5 should have reduced the influence of the order effect of the compatible and incompatible task, Nosek et al., 2005).

The prime type × target type × induction interaction was significant. With respect to the latencies in the EP paradigm, results showed that the prime type × target type × induction interaction was not further moderated by stimulus novelty, thereby providing evidence that associations of the concepts of ambiguity and clearness (and not only the stimuli representing these concepts) with positive or negative valence were established. Hütter and Rothermund (2020) noted that it is difficult to determine at what stage the generalization of evaluative learning takes place: in the learning phase or only later in the testing situation. Potentially, it could be that the learning phase only changed the mental representation of evaluations of specific stimuli and that generalization only occurred in the testing situation.

However, while the three-way interaction was significant (and did not depend on stimulus novelty), it was clearly different form predictions and showed almost the opposite pattern than expected. One post hoc explanation for this unexpected result (which partially shows a contrast effect) could be that participants had a nonconscious attempt to correct for the prime influence (see, Glaser, 2003). However, according to this account, contrast effects occur when there is a high motivation for accuracy and prime valence is extreme. Both conditions were not applicable to our study as we encouraged participants to be fast, and the prime valence, although manipulated, was probably not extreme. Another account explaining contrast effects is the psychophysical account (Klauer et al., 2009). According to this account, decision makers do not simply rely on the positive and negative activations of the target but also assess the increase or decrease of these activations relative to a recent starting point. As this starting point can vary in time, the evaluation window used for decision making can include (exclude) the activation of prime valence leading to assimilative (contrast) effects. According to this account, short (long) SOAs, speed (accuracy) focus, moderate (extreme) primes provide assimilative (contrast) effects by adopting the evaluation window. As the used SOA of 300 ms lies at the boundary of the short-long distinction it is possible that an exclusive evaluation window was adopted by participants leading to the found effects (which can be to some extend regarded as contrast effects; for example, the significant influence of prime type for positive targets for participants with the induction of associations of ambiguity with positive and clearness with negative showed longer latencies for ambiguous (vs. clear) primes preceding positive targets). Alternatively, the induction phase itself led participants to perceive the primes as misleading, resulting in a more exclusive evaluation window.

A limitation of the interpretation that associative learning took place in the sense that the concepts of ambiguity and clearness were associated with valence arises from the view that the response side (left or right) was instead associated with the concepts or together with the intended valence association during the induction phase. According to the response priming view this could cause or boost the observed prime type by target type by induction interaction. This explanation gets even more plausible if we consider evidence that the usual EP effect is driven by processes relying on response competition and not on facilitated processing of the target (Klauer, 1997; Klauer et al., 2005; Klinger et al., 2000; Voss et al., 2013). However, Ebert et al. (2009) provided evidence (via counterbalancing the allocation of attributes in the manipulation IAT) that both associative learning processes (valence and response side allocation) took place. This makes it unlikely that the effect we observed was purely driven by response priming based on learned allocations of concepts with left or right. In addition, the observed unexpected (partial contrast) effect could not be explained by learned response priming, when we consider the processes of the psychophysical account, as this account is based on changes in the activation of positive and negative *evaluative* information. We did not counterbalance the allocation of attributes in the manipulation IAT, as a switch in the positive and negative key assignment between the manipulation IAT and the EP paradigm would lead to additional asymmetrical task switching costs for the two different valence-counterbalanced manipulation IATs. This would make the results difficult to compare as different processes are probably involved. Furthermore, counterbalancing the valence dimension would result in different levels of difficulties between the valence-assignment-counterbalanced versions as it is known that for right-handed persons responding to positive (negative) affect is easier with the right (left) hand, whereas for left-handed persons this is reversed (de la Vega, Dudschig, De Filippis, Lachmair, & Kaup, 2013; Song, Chen, & Proctor, 2017). As we can assume that most of our sample was right-handed, introducing a manipulation-IAT with positive mapped on the left hand and negative on the right hand would have increased the difficulty for that version of the task for most participants. This may influence the associative learning process. For these reasons we refrained from counterbalancing the assignments of the valence labels in the manipulation IAT at the expense of clear interpretation of the results in terms of purely reflecting automatic activation of evaluation of ambiguity and clearness.

Reviewing the literature of TA revealed two approaches for manipulating the TA. Glover, Romero, Romero, and Petersen (1978) showed that a simulation game focused on cultural differences could impact the TA assessed with the Budner scale. Another attempt to change the directly reported TA was done by Endres, Camp, and Milner (2015) who investigated the influence of manipulated situational ambiguity on TA. However, as we used an indirect measure in order to investigate automatic activation of evaluation of ambiguity and clearness, an indirect approach based on associative learning seemed to be better suited to change the attitude towards the concepts than the aforementioned approaches for which an effect on direct measures was shown.

As the psychophysical account specifies conditions (e. g. the SOA) allowing for predictions when assimilative and when contrast effect should emerge, we manipulated the SOA in the next study in order to investigate whether the found partial contrast effect could be explained by this variable.

3.8 Study 6

This pre-registered study (see Appendix H) is a conceptual replication of study 5. This study investigated a post hoc explanation based on the psychophysical account (Klauer et al., 2009) for the interaction of prime type × target type × induced association found in study 5. As mentioned, this account provides an explanation why a standard EP effect may emerge for short SOAs and why a contrast effect (like the one found in study 5) may emerge for long SOAs. According to this theory, the used SOA of 300 ms in study 5 could be a long SOA producing contrast effects. In order to further investigate this interpretation, we experimentally varied the SOA (200 ms vs. 400 ms) between the blocks of the EP paradigm in this study. The sequence of SOA was counterbalanced.

3.8.1 Hypothesis

(1) We expected that the prime type × target type × induced association is further moderated by the SOA. A visualization of the predicted interaction for the short SOA (200 ms) of prime type × target type × induced association can be found in Figure 15 (study 5). For the long SOA (400 ms) the two prime type × target type interactions switch positions. A detailed textual description of this prediction can be found in the pre-registration (Appendix H).
(2) Furthermore, we expected that this four-way-interaction is stronger for stimuli used in the manipulation phase than for unknown stimuli.

3.8.2 Method

Participants. Data were collected for one week from April 29 to May 6, 2019. All participants were recruited via the mailing list of the University of Tübingen. We only include data from participants who finished the study completely. Only participants could take part in the study who actively confirmed that they speak German on native speaker level and that they did not use a smartphone for participating in the study. We collected full data sets from 332 participants (257 females, 70 males, 3 diverse, and 2 with missing gender information) with a mean age of 27.96 years (*SD* = 10.62 years). 154 (178) participants were randomly assigned to the attitude manipulation condition ambiguous = positive + clearness = negative (ambiguous = negative + clearness = positive). The stimulus set for the manipulation phase was counterbalanced. 170

(162) participants had stimulus set A (B). 164 participants were assigned to the sequence of an SOA of 200 ms in the first block and 400 ms in the second block and 168 participants had the reversed sequence. All participants participated in exchange for course credit or could take part in a lottery for one of ten 20 Euro vouchers for a chocolate shop located in Tübingen.

Design. We used a 2 (induced association: ambiguity = positive + clearness = negative vs. ambiguity = negative + clearness = positive) × 2 (prime type: ambiguous vs. unambiguous) × 2 (target type: positive vs. negative) × 2 (stimulus novelty: same as in induction phase vs. different from induction phase) × 2 (SOA: 200 ms vs. 400 ms) design with repeated measures on the last four factors. As in study 5, the specific stimulus set used for the manipulation phase was counterbalanced across participants.

Materials and Procedure. All stimuli were identical to those used in study 5. The procedure was identical with the exception that the manipulation check IAT was not included in this study.

3.8.3 Results

We applied a multilevel model. The same random effects were estimated as in previous studies. In the fixed effects model, we regressed the logarithmized reaction times in the EP paradigm on prime type × target type × induction × stimulus novelty × SOA and all subordinate four-way, three-way, two-way interactions and the main effects. The stimulus set (A or B) was included as a covariate. In order to bind error variance, we included a two-way interaction of the target type with the target type of the previous trial. The target type of the previous trial as a main effect was also included.

Like in the previous studies, we excluded all trials in the priming paradigms with reaction times below 300 ms (0.80 %) and above 3000 ms (0.24 %) and all false classifications (4.29 %) from the analysis. In total we excluded 4.90 % trials of the EP paradigm. This resulted in 60,623 remaining valid trials for the EP paradigm.

Random effect estimation for variance (*SD*) for the intercepts of participants was 0.0169 (0.1300), for the target stimuli 0.0008 (0.0274) and for the prime stimuli 0.0000 (0.0020). Fixed effects of the model can be seen in Table 27.

Table 27

Fixed effects of mixed-effect model (study 6).

	b	*SE b*	*df*	*t*	*p*
Intercept	6.4621	0.0091	122	707.68	< .001
Stim set	-0.0159	0.0057	21	-2.81	.01
Target type N-1	0.0052	0.0008	59594	6.20	< .001
Target type	0.0236	0.0057	21	4.17	< .001
Prime type	-0.0003	0.0009	21	-0.37	.71
Induction	0.0047	0.0072	330	0.65	.52
Stimulus novelty (SN)	-0.0069	0.0008	59576	-8.15	< .001
SOA	0.0199	0.0008	59579	23.53	< .001
Target type N-1 × Target type	0.0005	0.0008	59607	0.60	.55
Target type × Prime type	-0.0003	0.0008	59580	-0.40	.69
Prime type × Induction	0.0005	0.0008	59576	0.57	.57
Target type × Induction	0.0005	0.0008	59577	0.60	.55
Prime type × SN	-0.0005	0.0008	59593	-0.62	.53
Target type × SN	-0.0001	0.0008	59577	-0.11	.91
Induction × SN	0.0010	0.0008	59576	1.22	.22
Prime type × SOA	-0.0015	0.0008	59574	-1.81	.07
Target type × SOA	-0.0019	0.0008	59576	-2.30	.02
Induction × SOA	-0.0011	0.0008	59579	-1.29	.20
Stim new × SOA	0.0016	0.0008	59576	1.90	.06
Target type × Prime type × Induction	-0.0006	0.0008	59577	-0.75	.46
Target type × Prime type × SN	0.0006	0.0008	59576	0.70	.48
Prime type × Induction × SN	-0.0005	0.0008	48538	-0.57	.57
Target type × Induction × SN	0.0011	0.0008	59595	1.28	.20
Target type × Prime type × SOA	-0.0023	0.0008	59576	-2.76	< .01
Prime type × Induction × SOA	0.0002	0.0008	59574	0.27	.78
Target type × Induction × SOA	0.0010	0.0008	59575	1.20	.23
Prime type × SN × SOA	0.0007	0.0008	59577	0.87	.38
Target type × SN × SOA	-0.0002	0.0008	59577	-0.30	.77
Induction × SN × SOA	-0.0004	0.0008	59575	-0.43	.67
Target type × Prime type × Induction × SN	-0.0006	0.0008	59576	-0.69	.49
Target type × Prime type × Induction × SOA	-0.0006	0.0008	59575	-0.75	.45
Target type × Prime type × Stim new × SOA	-0.0002	0.0008	59576	-0.28	.78
Prime type × Induction × SN × SOA	0.0001	0.0008	59576	0.13	.90
Target type × Induction × SN × SOA	0.0002	0.0008	59577	0.22	.82
Target type × Prime type × Ind. × SN × SOA	-0.0002	0.0008	59577	-0.19	.85

Note. Effect coding was used for the dichotomous variables: Target type and Target type N-1 (-1 = positive, 1 = negative), Prime type (-1 = unambiguous, 1 = ambiguous), Stim Set (-1 = B, 1 = A Stimulus novelty (-1 = new stimuli, 1 = stimuli from induction phase), Induction (-1 = ambiguity was paired with positive and clearness was paired with negative, 1 = ambiguity was paired with negative and clearness was paired with positive), SOA (-1 = 400 ms, 1 = 200 ms). This analysis was based on 59980 observations from 332 participants.

As can be seen in Table 27, the main effects of target type and target type N-1 were similar to the previous studies. Like in Study 5, there was a main effect of Stimulus novelty. Responses for stimuli known from the induction phase were faster than responses for new stimuli. The main effect of SOA showed that the responses for the 400 ms (vs. 200 ms) SOA were faster. There neither was a prime type × target type × induced association × SOA interaction (H1) nor was this four-way interaction further moderated by the stimulus novelty (H2). Instead, there was a significant prime type × target type × SOA interaction. Figures 18 and 19 visualize this interaction from different perspectives.

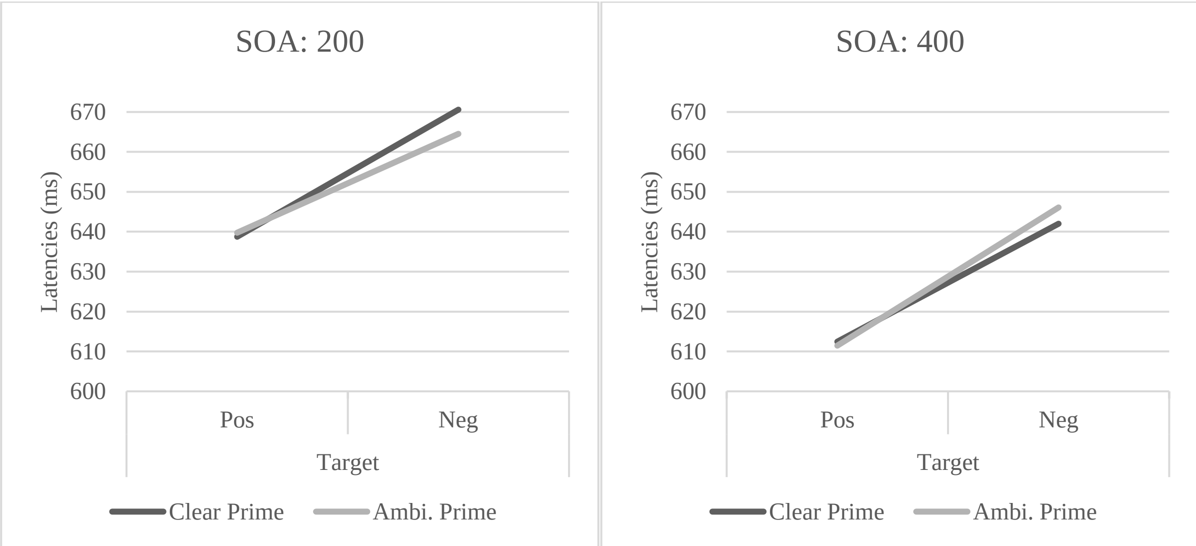


Figure 18. Visualization of the three-way interaction of prime type, target type and SOA. Left (Right) plot shows the predicted values of the prime type × target type interaction for the short (long) SOA. As the relevant comparison were within-subject comparisons, we did not include standard errors of the mean.

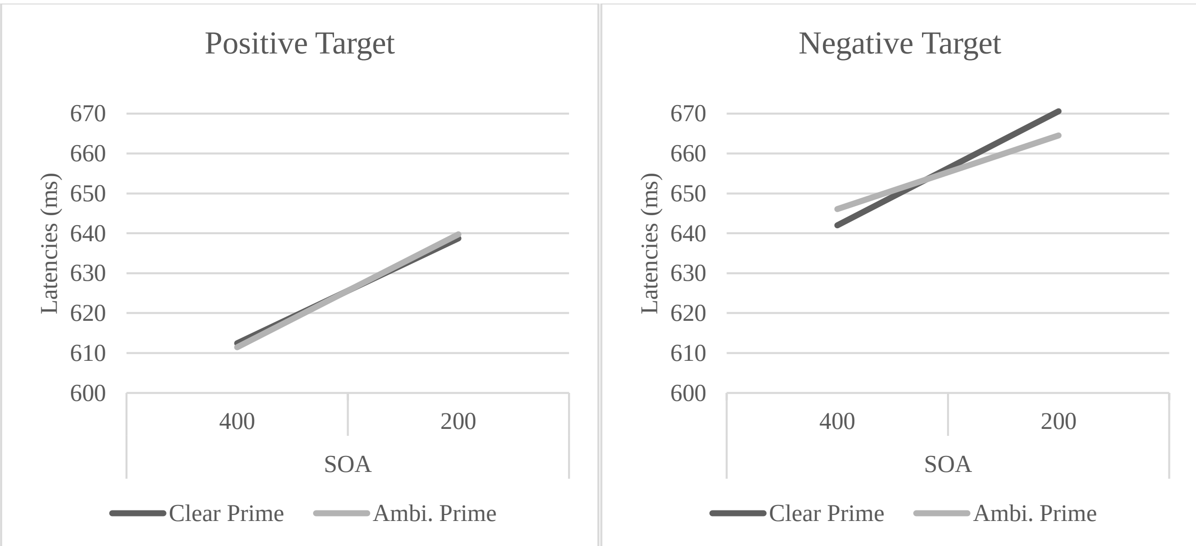


Figure 19. Visualization of the three-way interaction of prime type, target type and SOA. Left (Right) plot shows the predicted values of the prime type × SOA interaction for positive (negative) Targets. As the relevant comparison were within-subject comparisons, we did not include standard errors of the mean.

Dummy coding SOA revealed that the prime type × target type interaction was significant for the 200 ms SOA, $b = -0.0027$, $SE\ b = 0.0012$, $t = -2.24$, $p = .03$, and marginally significant for the 400 ms SOA, $b = 0.0020$, $SE\ b = 0.0012$, $t = 1.67$, $p = .096$. The three-way interaction was mainly driven by the SOA dependent influence of prime type for negative targets but not for positive targets. Figure 19 shows the three-way interaction separated by target type. Dummy coding target type revealed that the prime type × SOA interaction was significant for negative targets ($b = -0.0039$, $SE\ b = 0.0012$, $t = -3.24$, $p < .01$) but not for positive targets ($b = 0.0008$, $SE\ b = 0.0012$, $t = 0.67$, $p = .50$). The simple main effect of prime type for negative targets for a SOA of 200 ms was significant, $b = -0.0045$, $SE\ b = 0.0017$, $t = -2.62$, $p < .01$. The latencies were longer for clear (vs. ambiguous) primes preceding a negative target. For the 400 ms SOA, the simple main effect of prime type for negative targets was marginally significant, $b = 0.0032$, $SE\ b = 0.0017$, $t = 1.83$, $p = .07$. Descriptively, latencies were shorter for clear (vs. ambiguous) primes preceding a negative target.

Post-hoc analysis. In order to interpret the prime type × target type × SOA interaction we analyzed the data from the induction phase. If there was evidence that the sample on average had a strong association of ambiguity with negative valence and clearness with positive valence, it would provide an interpretation of this three-way interaction. We used a multilevel model with participants and stimuli as random effects. In the fixed effects model, we regressed the logarithmized reaction times of the third block of the IAT on induction × 3 different contrast variables designed to compare the 4 different stimuli categories (ambiguous, clear, positive and negative). The first contrast variable was coded in a way that ambiguous and clear stimuli were contrasted with positive and negative stimuli (0.5, 0.5, -0.5, -0.5). The second contrast variable was coded to contrast the ambiguous and clear stimuli (-0.5, 0.5, 0, 0), and the last contrast variable was coded in order to compare positive and negative stimuli (0, 0, 0,5, -0,5). The stimulus set (A or B) was included as a covariate. We excluded all trials of this block with reaction times below 300 ms (0.54 %) and above 3000 ms (3.11 %) and all false classifications (8.50 %) from the analysis. In total we excluded 11.31 % trials of this block. This resulted in 14,133 remaining valid trials of the third block of the induction phase. Random effect estimation for variance (*SD*) for the intercepts of participants was 0.0419 (0.2046) and for the stimuli 0.0023 (0.0477). Fixed effects of the model can be seen in Table 28. Visualization of the of the interaction of stimulus type with induction can be seen in Figure 20.

Table 28

Fixed effects of mixed-effect model of latencies in block 3 (Study 6).

	b	*SE b*	*df*	*t*	*p*
Intercept	6.7810	0.0186	100	365.12	< .001
Stim Set	-0.0185	0.0135	262	-1.37	.17
Induction	-0.0260	0.0125	425	-2.08	.04
Contrast 1	0.2162	0.0210	43	10.30	< .001
Contrast 2	0.0262	0.0211	44	1.24	.22
Contrast 3	-0.0588	0.0209	42	-2.82	< .01
Induction × Contrast 1	-0.0266	0.0078	13791	-3.40	< .001
Induction × Contrast 2	-0.0028	0.0081	13793	-0.34	.73
Induction × Contrast 3	0.0015	0.0075	13780	0.20	.84

Note. Effect coding was used for Induction (-1 = ambiguity + positive and clearness + negative, 1 = ambiguity + negative and clearness + positive), Stim Set (-1 = B, 1 = A). The four stimuli types: ambiguous, clear, positive, negative were represented by the three contrast variables (see text).

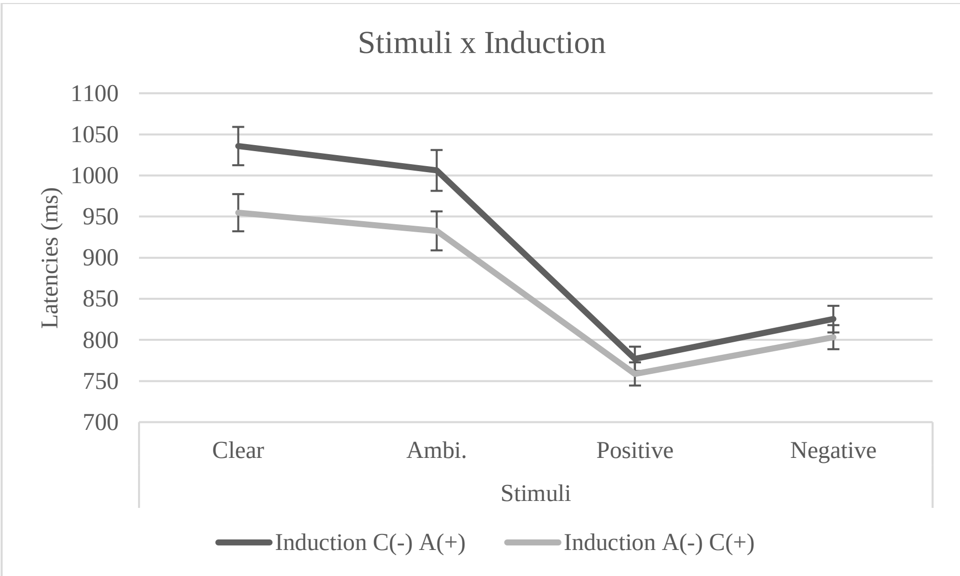


Figure 20. Predicted values of latencies as a function of stimulus type and induction. C(-) A(+) means clearness was paired with negative (location left in block 3) and ambiguity was paired with positive (location right in block 3). A(-) C(+) means ambiguity was paired with negative (location left in block 3) and clearness was paired with positive (location right in block 3). Error bars represent standard errors of the mean.

The significant main effect of contrast variable 3 showed that (similar to the responses to the targets in the EP paradigm) the responses for words with positive valence were faster than the responses for negative valence. More importantly, the significant main effect of induction showed that the responses for the condition ambiguity paired with negative and clearness paired with positive were faster than the responses for the condition clearness paired with negative and ambiguity paired with positive. This main effect is further qualified by the significant interaction with the contrast variable 1, indicating that this effect was stronger for clear and ambiguous stimulus than for positive and negative stimuli. This can be taken as evidence that the pairing of clearness with negative and ambiguity with positive is more difficult than the other pairing. This can be interpreted as indicating that the sample in general had a stronger association of ambiguity with negative valence and clearness with positive valence than ambiguity with positive valence and clearness with negative valence.

To rule out the possibility that the two groups constituted by the variable induction differed in their response times for clear and ambiguous stimuli per se, we additionally analyzed the first block of the induction phase in which ambiguous and clear stimuli were categorized. Again, we used a multilevel model with participants and stimuli as random effects. In the fixed effects model, we regressed the logarithmized reaction times of this block of the induction phase on induction × stimulus type (ambiguous, clear). The stimulus set (A or B) was included as a covariate. We excluded all trials of this block with reaction times below 300 ms (0.30 %) and above 3000 ms (9.25 %) and all false classifications (16.60 %) from the analysis. In total we excluded 23.57 % of the trials of this block. This resulted in 6,090 remaining valid trials of the first block of the induction phase. Random effect estimation for variance (*SD*) for the intercepts of participants was 0.0549 (0.2343) and for the stimuli 0.0039 (0.0624). Fixed effects of the model can be seen in Table 29.

Table 29

Fixed effects of mixed-effect model for latencies in Block 1 (study 6)

	b	*SE b*	*df*	*t*	*p*
Intercept	7.1060	0.0187	73	380.39	< .001
Stimulus Set	-0.0133	0.0187	73	-0.71	.48
Induction	-0.0134	0.0137	321	-0.94	.33
Stimulus type	-0.0072	0.0135	21	-0.54	.60
Induction × Stim type	0.0066	0.0044	5770	1.51	.13

Note. Effect coding was used for Induction (-1 = clearness (left) and ambiguity (right), 1 = ambiguity (left) and clearness (right)), Stimulus Set (-1 = B, 1 = A), Stimulus type (-1 = ambiguity, 1 = clear).

As can be seen in Table 29 the two induction groups did not differ in their response times. This can be taken as evidence that the localization of ambiguity and clearness (left or right) does not play a role and that the response time difference in block 3 between the two induction conditions reflected a stronger association of ambiguity with negative and clearness with positive valence than the reversed pairing.

3.8.4 Discussion

This study investigated whether the unexpected contrast effect found in study 5 can be explained by the variable SOA. Our first hypothesis was that the prime type × target type × induction interaction was further moderated by the SOA. Again, as in study 5 we counterbalanced the stimulus set used for the induction phase in order to investigate if the potential learning of associations is bound to specific stimuli or generalized to concepts. Our second hypothesis was that the effects in the EP paradigm should be stronger for known stimuli compared to new stimuli. However, we found no support for both hypotheses. Instead our data showed that the prime type × target type interaction was moderated by the SOA. Assuming that the sample in general had a stronger association of ambiguity (clearness) with negative (positive) valence compared to the association of clearness (ambiguity) with negative (positive) valence, the observed data pattern in the EP paradigm makes sense: assimilation effect at short SOA and contrast effect at long SOA. Support for this interpretation comes from the analysis of the reaction times of the third block of the IAT which showed that the mean reaction times were faster for participants which received the pairing ambiguity (clearness) with negative (positive) compared to participants which had the reversed pairing, for both clear and ambiguous stimuli. Additional analyses of the first block of the IAT revealed that the two groups constituted by the side assignments of ambiguity and clearness did not differ in their reaction times for ambiguous and clear words per se. These post hoc analyses of the manipulation IAT as well as all previous studies (via the mean IAT *D* score or the results of the EP paradigm or the less effective change of attitudes in the direction ambiguity = positive/clearness = negative) provide evidence that there is a stronger association of ambiguity with negative and clearness with positive in general.

4 General Discussion

The present work investigated whether an automatic activation of evaluations of having one vs. two meanings in mind occurs and whether this activation of evaluations is interindividually different in terms of valence associated with these two different mental states. To put the question in a nutshell, do some people have a fast and unintentional positive evaluation of the core of the phenomenon ambiguity (having more than one meaning in mind) while others have a negative one? In order to investigate this question, we conducted five studies using the EP paradigm in which ambiguous and clear words were used as primes and words with positive and negative valence served as targets. In three of these studies, we additionally assessed the attitude toward ambiguity and clearness in relation to each other via the IAT. In the other two studies, we tried to induce associations of opposite valence with ambiguity and clearness prior to the EP paradigm.

Study 2 provided evidence for interindividual differences in automatic activation of evaluations of learned ambiguous and clear words. However, there was only an interaction of prime type and target type for one *SD* above the *D* score of zero (for which the IAT indicated that clearness is more liked than ambiguity/ambiguity more disliked than clearness). It showed, as predicted, that ambiguous (vs. clear) primes preceding positive targets led to longer latencies. However, there was no influence of prime type for negative targets. As already discussed (see section 3.4.4), the learning paradigm prior to the EP paradigm could have established associations of the words representing the categories (ambiguity and clearness) with exactly these categories in addition to the pre-activation of associations with the meaning(s). Therefore, the EP paradigm in this study did not allow to distinguish between unintentional activation of evaluations of the mental state of having one vs. two meanings in mind and unintentional activation of evaluations of the superordinate concepts of ambiguity and clearness. In this sense, although there is evidence that this task is mainly driven by individual stimulus properties (Livingston & Brewer, 2002), the EP paradigm with the learning paradigm (and probably established associations with the superordinate categories) could be more similar to the IAT for which there is evidence that it measures to a stronger extent the automatic associations of the concepts (instead of the individual stimuli) with valence (De Houwer, 2001, 2003b, 2008; Nosek et al., 2007).

To further strengthen the evidence of interindividual differences in the automatic activation of evaluations of having one vs. two meanings in mind, we added a condition without a learning phase prior to the EP paradigm in study 3. Indeed, this condition provided evidence for this effect: there was a significant interaction of prime type, target type and IAT *D* score. However, the predicted latencies as a function of prime type and target type for participants liking ambiguity more than clearness and for participants liking clearness more than ambiguity (see Figure 13) showed the expected pattern only when the sample mean was considered as neutral point, (cf. visualization of hypothesis in Figure 8). Both two-way interactions of prime type and target type were not significant, however. By showing that the three-way interaction of prime type, target type, and IAT *D* score was not further moderated by the variable capturing the information of whether the same or different stimuli were used in the EP paradigm and the IAT, the evidence for the evaluative character of the mental state of having one vs. two meanings in mind was strengthened. However, potentially due to procedural changes between the studies, the findings for the condition with a learning phase were not replicated.

Unfortunately, we could not replicate the interaction of prime type, target type, and IAT *D* for the condition without a learning phase (study 4). Nevertheless, there was a strong prime type by target type interaction found for the sample mean in that study. The influence of prime type for positive and for negative targets was significant. It showed that ambiguous compared to clear primes lead to shorter (longer) latencies for negative (positive) targets. This provided evidence that, on average, ambiguous primes were automatically evaluated as more negative (or less positive) than clear primes.

By experimentally manipulating the attitude towards ambiguity and clearness in study 5 and 6, we tried to counteract potentially insufficient construct related variance. In study 5, the manipulation of the attitude towards ambiguity and clearness did indeed significantly interact with prime type and target type, but opposite to the direction expected. The prime type and target type interaction was only significant for participants for whom the association of ambiguity with positive and clearness with negative was induced and only the influence of prime type for positive targets was significant (latencies for positive targets were longer when ambiguous (vs. clear) primes preceded the targets). This unexpected partial contrast effect was discussed with reference to the psychophysical account (Klauer et al., 2009).

Therefore, we further investigated this post-hoc explanation by manipulating the SOA in study 6. However, the manipulation of the attitude towards ambiguity and clearness had no influence on latencies in the EP paradigm in this study. However, there was a significant prime type × target type × SOA interaction. Only for negative targets there was an interaction of prime type and SOA. It showed that for this target type ambiguous compared to clear primes led to shorter (longer) latencies for a short (long) SOA. A closer look at the attitude induction phase provided evidence for a stronger association of ambiguity with negative valence and clearness with positive valence which is in line with the IAT results from studies 1, 3, 4 and with the interpretation of data of the manipulation check IAT of study 5. A stronger association of ambiguity with negative valence and clearness with positive valence would perfectly explain the found result of the moderated prime type × target type interaction by the factor SOA in terms of contrast (long SOA) and assimilation (short SOA) effects.

4.1 Interpretation of Results

Integrating all results, we have to admit that we found only tentative evidence for *interindividual* differences in automatic activation of evaluations of the concepts of ambiguity vs clearness (evidence only in study 2), and in particular of the mental states characterized by activation of one vs. two meanings of words (evidence only in study 3). However, data from study 2, 4 and indirectly from study 6 indicate that, on average, the automatic evaluations activated by the concept of ambiguity and the mental state of having multiple meanings in mind are more negative (or less positive) compared to the automatically activated evaluations of the concept of clearness and having only one meaning in mind. This is interesting insofar as the ambiguity we presented in the EP paradigm was context-free (we only provided context in the learning phase of study 2 via sentences in order to pre-activate both meanings) and no choice between the meanings had to be made. For ambivalence (the state of having negative and positive evaluations activated derived from one attitude object) it was shown that negative affect only occurred when an evaluative decision had to be made (van Harreveld, Rutjens, Rotteveel, Nordgren, & van der Pligt, 2009) and the conflicting evaluative information of the attitude object is relevant in the context in which the decision occurs (Nohlen, van Harreveld, Rotteveel, Barends, & Larsen, 2016). Therefore, for ambiguity it can be speculated that an unintentional fast process to select the "correct" meaning takes place, which leads to a more negative evaluation of ambiguity when there is no contextual information for disambiguation.

We often used the terminology of *automatic activation of evaluation of ambiguity*. We already discussed what is meant by ambiguity (concept vs. mental state). However, we think that it is necessary to discuss what the other terms "automaticity", "activation", and "evaluation" imply as this information is relevant for the interpretation of the findings.

Activation and Evaluation. The question arises what we mean by activation and what is the subject of activation. Obviously, the term activation implies that there is something that gets activated. On the one hand, according to Fazio's view that attitudes are object-evaluation associations of varying strength (Fazio, 1990, 2001, 2007; Fazio et al., 1986), the subject of automatic activation can be the evaluation proper, as within this framework it is assumed that an evaluative summary can be directly associated with the attitude object. According to this theory no further mediating step is needed. On the other hand, as evaluation can be considered as the *process* of identifying the liking of an object (De Houwer, 2009) one could also consider the case that the result of this process is not activated automatically via associative links, but needs prior activation of information which then starts the process of evaluation. In that sense the terminology of activation of evaluation would imply that there already is some (evaluative, useful) information stored in mind which is activated. In other words the attitude object (in a specific context) activates a “relatively stable set of representations of a stimulus” (William A. Cunningham & Zelazo, 2007, p. 97) which builds the basis of the evaluative process. It needs to be mentioned that the specification of the underlying process and the kind of representation of evaluation, for example associative or propositional (for a recent overview of evidence for the importance of propositional representations in evaluation see, De Houwer, Van Dessel, & Moran, 2020), is irrelevant for our main questions. Namely, whether ambiguity (concept and/or mental state) could be the subject of an automatically (unintentionally and fast) activated evaluative process or an automatically activated evaluative summary and whether the result of this evaluative process or the valence of the activated summary evaluation is interindividually different such that there is a substantial variation in liking. In our view, speaking of “automatic activation of evaluation” is neutral in terms of possible underlying mediating mechanisms, but implies that there already exists evaluatively useful information or a summary evaluation that gets activated. Furthermore, this terminology does not make any assumption about the quality of evaluation in terms of which type of valence is activated (see Itkes et al., 2017, for empirical support for the distiction between affective valence and semantic valence). Does ambiguity trigger a (diffuse) “hot” affective feeling (affective valence) or is some kind of “cold” knowledge about the valence (semantic valence) of ambiguity involved? Our paradigm does not allow to make any statements about this.

Automaticity. In our view applying an EP paradigm without a prior learning or manipulation phase permits the strongest claim about the automaticity of activation of evaluation of having one vs. two meanings in mind. However, which features of automaticity (see, Bargh, 1994; Moors, 2016) are likely involved in the EP paradigm using lexically ambiguous and clear words as primes and valent words as targets? By looking at different features of automaticity which specify under which conditions a process can occur (operating conditions), we will not be making any assumptions about the operating principles specifying which process underlies the evaluation (for information about the relevance of the distiction between processes and conditions see, Gawronski & Bodenhausen, 2009, 2014).

As the process of evaluation in the present paradigm is based on the automatic activation of meaning(s), we will shortly describe in which sense this process can be considered as automatic. The meaning(s) of a word is activated fast, efficient (it requires only the direction of attention to the word) and uncontrollable in both senses of controllability: the process of meaning activation starts without having a goal to do so and the process cannot be suppressed volitionally. This is demonstrated by the highly robust effect found in the Stroop task (Stroop, 1935): the activation of the meaning of the color word starts unintentionally and cannot be suppressed although it is task irrelevant (Augustinova & Ferrand, 2014). As meaning activation is automatic in the mentioned sense, assumed discrete mental states characterized by having one vs two (or many) meanings of an isolated, visually presented word in mind can be assumed to occur automatically in the same way (see chapter 2.2 for an overview of studies providing evidence for automatic meaning activation of lexical ambiguous words). Now we will have a closer look at features of automaticity in reference to the process of evaluation in the EP paradigm.

Efficiency. The efficiency feature of automaticity (concerning the question of whether the process needs attention?) is difficult to determine in reference to the process of evaluation of ambiguity in the EP paradigm. On the one hand, there is evidence that the process of evaluation requires little attentional capacity, as for example there is evidence that an EP effect can be observed even for subliminally presented primes (e.g., Gibbons, 2009). While this points to the quantitative dimension of attention, the feature-specific attention allocation model (Spruyt et al., 2012; Spruyt et al., 2007; Spruyt et al., 2009) postulates that the amount of semantic activation (including activation of evaluative information) depends on the assignment of attention to a specific stimulus dimension (qualitative dimension of attention). In the standard EP paradigm, the attention to the evaluative dimension is provided by the task (evaluative decision of the targets). As we did not directly manipulate the allocation of attention to the ambiguity dimension, it remains unclear if this could play a role for our results. Furthermore, as there are two sequential processes (the *evaluation* and the *translation* of the result of this evaluation process into observable responses) taking place (Gawronski, Cunningham, LeBel, & Deutsch, 2010; Moors, Spruyt, & De Houwer, 2010), in the EP paradigm like in all indirect tasks (aiming to assess an evaluation) it is even more difficult to make statements about the efficiency feature of automaticity regarding the first step: the evaluation of the prime.

Consciousness. Due to not having a learning phase (which word is ambiguous or not) or an attitude induction phase prior to the EP paradigm, attention to the distinctive feature of ambiguity is probably very weak. If there was evidence that the participants did not consciously recognize the distinctive feature (at least for short SOAs), the automaticity feature of unconsciousness could be applied to this state of having multiple meanings in mind triggered by the lexical ambiguous words, too. Of course, even if semantic meaning activation is unconscious this does not necessary imply that we do not have conscious awareness of the evaluation (for further examples when people are conscious about the evaluation but not of the reason of evaluation, see Gawronski & Bodenhausen, 2014). However, as our paradigm was not designed to provide information about the consciousness of semantic activation nor about the consciousness of evaluation triggered by the mental state of having multiple meanings in mind, we do not make any claims about it.

Control. Without a strong top-down influence, the process of evaluation in the EP paradigm can be considered uncontrolled in the promoting sense: there is no goal to start the process, it happens unintentionally. However, in the counteracting sense of control, an evaluative priming effect can be shaped by strategic top-down influence (Alexopoulos et al., 2017). However, as mentioned, the effect in the EP paradigm (like all tasks based on irrelevant features) is based on two steps: processing of the valence of the prime (evaluation) and a translation of the first step into what we can observe, here, the influence of the evaluation on the target classification (Gawronski et al., 2010; Moors et al., 2010). It is unclear whether the controllability of the EP effect is due to the inhibition of valence activation or the inhibition of consequences of this activated valence on the target classification (e.g., due to segmentation of the information stream derived from prime and target). Therefore, it could be that the activation of valence is to some extend uncontrolled (in the counteracting sense), whereas the effects on the target classification can be controlled.

Time. The process of evaluation (and the process of translation into observable responses) can be considered fast as EP effects are more reliable (Klauer et al., 2009) and more pronounced (Herring et al., 2013) for short SOAs (below 300 ms).

Although we already discussed potential reasons why direct and indirect measures did not correlate for the construct attitude towards ambiguity in section 3.3.4, we will now look at this topic from a different perspective, potentially explaining the inconsistent finding of moderation of the EP effect by the IAT score, too. Of course, by introducing this perspective, we cannot rule out that methodological reasons (e.g. low reliabilities) or the influence of construct unrelated processes (e.g. recoding processes) or explanations derived from dual-process theories (e.g. arguing that the automatic evaluative reaction to ambiguity is regarded as invalid for an evaluative judgment according to the framework of the Associative-Propositional Evaluation model (Gawronski & Bodenhausen, 2006, 2014) were nevertheless relevant determinants that prevented finding a correlation between direct and indirect measures. We already introduced the distinction between the phenomenon of ambiguity in general and the specific mental state of having multiple distinct interpretations in mind (subjective ambiguity) in section 2.3. It could be that direct and indirect measures of the attitude towards ambiguity have no relation as they virtually assess the attitude towards two different entities. In our view, direct measures of attitude towards ambiguity/intolerance of ambiguity usually assess the evaluative judgment of the abstract phenomenon. Indirect measures like the EP paradigm, especially when using lexical ambiguity as stimulus, probably assess the evaluation of the mental state of having multiple distinct meanings in mind.

It is obvious that the evaluations of different entities could be unrelated. For example, the intellectual play with two or more meanings which is probably associated with the phenomenon of ambiguity could be considered as pleasant or interesting even though the state of having multiple meanings in mind could cause a negative feeling. In a similar vein, we can evaluate a sad movie positively despite it causing negative affect. Therefore, ambiguity could be a special attitude object in the sense that the attitude towards the phenomenon of ambiguity and the attitude towards the mental state of having multiple meanings or interpretations in mind can be quite different. Therefore, to the extent that direct and indirect measures assess evaluations of different attitude objects, this could explain the finding of uncorrelatedness.

The same argument (assessment of evaluations of different entities) could also be made in order to explain the inconsistent moderation of the EP effect by the IAT. We already mentioned that the IAT likely acts more on a conceptual level and is less driven by stimulus irrelevant (e.g. valence) properties than the EP paradigm. In this sense, only the EP paradigm allows to draw conclusions for the unintentional activation of evaluations of the mental state of having one vs two meanings in mind. The IAT on the other hand does not allow to draw conclusions from this perspective, as in this task, the targets need to be classified explicitly based on the discriminative feature. Therefore, using the IAT as a moderator of the EP effect of course may not be the theoretically optimal choice. The difference of what the IAT and the EP paradigm measure could potentially explain a non-significant three-way interaction although the possibility remains that both the automatic evaluations of the abstract concept of ambiguity and the mental state triggered by ambiguous stimuli differ interindividually. From this point of view, it would seem better to use two indirect measures which are both equally driven by irrelevant stimulus properties. However, compared to the IAT, other available indirect measures, which are probably more similar on a process level, are less reliable (Bar-Anan & Nosek, 2014). Furthermore, using a more similar indirect measure would probably lead to an increase of method-specific covariance which would complicate the interpretation of results as well.

Furthermore, the goal of measuring the attitude towards the same object across participants and the same object via different measures is further challenged by recent research highlighting the propositional (in this case especially relational) perspective of representations relevant for evaluation (De Houwer et al., 2020). They argue that as indirect tasks like the IAT do not specify the relation of, for example, attribute concepts and target concepts, there could be interindividual variation in the specification of the relation. For example, the pairing of ambiguity and positive could be specified by the propositions: “ambiguity is positive” or “ambiguity occurs in positive contexts” or “ambiguity is predictive of positive situations”, which would result in different IAT scores although the evaluation of the concept of ambiguity could potentially have no interindividual variation. So, different applied relations of concepts could reduce the validity of the IAT (the same is true for the relation of prime and target in the EP paradigm).

4.2 Alternative explanations

One could argue that the evidence for the predominantly negative evaluation of lexically ambiguous words compared to unambiguous words is not due to a pre-existing negative attitude towards the mental state of having two meanings in mind but due to the experience of influent processing which causes a negative affect and which influences the stimulus evaluation (for an overview of how affective reaction to ambiguity is mediated by fluency see, Halberstadt & Winkielman, 2013). However, Owen, Halberstadt, Carr, and Winkielman (2016) showed that gender-ambiguous faces were evaluated more negatively compared to gender-unambiguous faces only when participants had to classify gender prior to the evaluative rating phase. When the faces were classified on another dimension, the ambiguous and unambiguous faces did not differ in their evaluation. Furthermore, they showed that the fluency (as indicated by time needed to classify a face) fully mediated the evaluation of the faces. Their results are important insofar as one could draw the conclusion that processing fluency of ambiguous stimuli is not higher per se but could depend on directing attention to and asking for classification of other non-ambiguous features of the stimuli. In our paradigm, we did not request a selection of one meaning of a lexically ambiguous word. Therefore, the effort of processing lexically ambiguous words is probably not higher than that of processing lexically unambiguous words. However, one could argue that learning which word is ambiguous and which word is unambiguous prior to the EP paradigm (in studies 2 and 3) indeed directed attention to the discriminant feature and probably resulted in the desire to select the intended meaning of the meanings of the ambiguous words. This would probably decrease the processing fluency, which in turn could have led to the evidence for the generally (i.e., irrespective of measured associative strength via the IAT) more negative evaluation of ambiguous words compared to unambiguous words (see results of study 2). However, it should be mentioned that this account could also predict the opposite pattern if participants tried to distinguish between lexically ambiguous and unambiguous words. As the search for meaning(s) is terminated earlier for lexically ambiguous words than for lexically unambiguous words, for which an exhaustive search is discussed (Forster & Bednall, 1976), one could argue that the processing of unambiguous words is less fluent than that of ambiguous words. This would imply that unambiguous words trigger a more negative evaluation than ambiguous words, which does not fit our findings. Moreover, the assumed asymmetrical fluency of ambiguous and unambiguous words could have an impact on the translation of the evaluation of the prime into observable responses (besides the potentially triggered evaluation based on that feature) as well. Alexopoulos et al. (2012), via manipulating the fluency of the primes (gradual demasking and color contrast), showed that fluent primes compared to disfluent primes led to reduced priming effects This was interpreted by the authors in reference to the segmentation hypothesis (Fiedler et al., 2011). Fluent primes lead to a completed mental episode which constitutes a segment boundary, functionally separating prime information from target information. Applying this logic to our assumed asymmetry of fluency of ambiguous and unambiguous words would mean that it is more difficult to measure automatic evaluations of ambiguous compared to unambiguous words via the EP paradigm.

While the fluency account does not argue against evaluative responding per se, but only highlights the role of processing fluency for the evaluation, the next alternative explanation questions that the observed responses are indicative of an evaluation in general. Ambiguous and clear words could prime a response associated with the number of meanings not based on valence. Dehaene, Bossini, and Giraux (1993) showed that larger (smaller) numbers were faster categorized with the right (left) hand in an odd-even judgment task. This Spatial-Numerical Association of Response Codes (SNARC) effect was shown to be robust across different tasks, stimuli and populations (Wood, Willmes, Nuerk, & Fischer, 2008). Even for tasks for which the numerical information is irrelevant for the response (e. g. color, orientation judgments) a

SNARC effect was observed (T. Mitchell, Bull, & Cleland, 2012). In most paradigms, numbers ranging from 1 to 9 were used. In contrast, in our tasks only words with one or two meanings were used, providing only numerical information for 1 and 2. If this dichotomy resulted in a stronger contrast between the two numbers, ambiguous words with two meanings were potentially more strongly associated with the right hand as they are the larger ones in our paradigm and clear words with one meaning were potentially associated more strongly with the left hand as they represent the smaller value. This association of space and numbers would predict a prime type × target type interaction in the EP paradigm in a way that ambiguous (clear) primes facilitate reactions to positive (negative) targets, as the valence label assignments were always negative (left) and positive (right). Across studies, we found evidence for the opposite pattern on a group level (disregarding the individual attitude towards ambiguity). Therefore, potential priming effects due to an association of space and number (of meanings) do not seem to play a role in our experiments.

We already mentioned that the two indirect measures (EP and IAT) as well as the investigation of the relation of both measures can only detect relative differences between the concepts. If there was interindividual variance of valence affecting both concepts in the same way (both positive or both negative) and hence no intraindividual evaluative differences between the concepts (disregarding non-construct related variance), the IAT *D* score would be zero and a prime type × target type interaction in the EP paradigm would be missing as well as a moderated interaction by the IAT *D* score. Therefore, theoretically, a non-significant three-way interaction of prime type, target type, and IAT *D* score does not necessarily mean that there are no interindividual attitudinal differences regarding the concepts of ambiguity and clearness.

However, the theoretical case that the evaluations of having one and two meanings in mind are nearly identical allowing for the non-significant three-way interaction (despite between subject variance of valence of both evaluation objects existing in the same way) seems unlikely as one could pose the question if an evaluation of *one* activated meaning would be adaptive in any way for the organism, as one activated meaning (although lexical ambiguity is a common phenomenon) constitutes the default state and lexical ambiguity is the exception, for which any (evaluative) reaction seems more plausible as potentially being adaptive. Therefore, unrelated evaluations of the state of having one and two meanings in mind seems more plausible than identical evaluations. A neutral or non-existent evaluation of one entity (probably clearness) and a non-neutral evaluation of another entity would result in a relative difference of evaluations between the entities which (if the evaluation of the non-neutrally evaluated entity is intraindividually stable and interindividually different) should be detected by the combined analysis of the two indirect measures employed here (i.e., IAT as moderator for the EP effect). Empirical results from correlational studies using direct measures (see 3.2) clearly speak for a negative relationship between the attitudes towards the two concepts. However, it must be noted that we never found a substantial relationship between direct and indirect measures of attitudes making the transfer of this relationship to indirect measures questionable. We already discussed the unrelatedness of direct and indirect measures in terms of the different attitude objects that the measures assess (see 4.1). Therefore, the negative relationship between ambiguity and clearness found using direct measures may not be indicative of the relationship between the attitude towards having one vs. two meanings in mind. In that domain, an unrelatedness due to a probably non-existent attitude towards the state of having one meaning in mind seems more plausible. However, during the EP paradigm (and the IAT as well), words with one meaning could acquire the opposite valence of words with two meanings as they were compared to them. For the IAT, it was shown that previously neutral objects showed an evaluative shift likely due to analogical learning principles (Hussey & De Houwer, 2018). For the EP paradigm, a symbolic evaluative generalization (see, Hughes et al., 2018) could take place in a way that people learn that one meaning words are the opposite of two meaning words. Together with the assumed pre-existing attitude towards two meaning words, people could infer that one meaning words possess the opposite valence. The evaluative generalization can be seen to depend on symbolic relations as the one and two meaning words are not related on other perceptual dimensions.

4.3 Limitations and future research

In all experiments, the word selection has been based on normative data of the pretest. Especially the identification of ambiguous words with two balanced meanings based on the normative data does not allow to draw the conclusion that the meanings are balanced for a specific person. If one uses lexically ambiguous words as stimuli, for which a lot of evidence exists that in general (see Chapter 2.2) both meanings get activated fast, efficiently and uncontrollably, it might still be the best way to select the normatively balanced words, as idiosyncratic selection of words with balanced meanings is difficult to impossible since one meaning is always faster accessible at a conscious level. However, regarding the serious limitation that we did not know if the two meanings of a lexically ambiguous word were balanced (in the sense that they are equally strongly activated at the same time) for a specific person based on the given normative data, one interesting option for future research would be to experimentally induce the number of meanings of items (e.g. pseudowords) and use these items as primes in the EP paradigm and as stimuli in the IAT. On the one hand, this would allow to create ambiguous and unambiguous items perfectly matched for several relevant dimensions. On the other hand, this would obviously result in more attention to the distinctive feature ambiguity for the primes in the EP paradigm which may counteract the EP effect.

The role of attention to the relevant prime feature for the EP effect is difficult, as on the one hand there is evidence that direct attention (e.g. via counting prime features) to the prime feature of interest (for multiply categorizable primes) is necessary for the EP effect of a specific feature (e.g., Gawronski et al., 2010; Olson & Fazio, 2003). On the other hand, according to the segmentation hypothesis (Alexopoulos et al., 2017; Fiedler et al., 2011) responding to specific prime features like, for example, grammatical gender (Fiedler et al., 2011; Exp. 1c) could result in a functional segmentation of the stream of information (prime and target constitute two separate mental episodes), resulting in an exclusion of the prime information from the target information and thus eliminating the standard EP effect. While these studies investigated the role of attention alignment to specific features of multiple categorizable primes, Simmons and Prentice (2006) investigated whether attention (attending to the prime word or ignoring the prime word by directing attention to a competing prime stimulus, which was a digital number) in general modulated the EP effect. They found that participants attending to the prime word had stronger EP effects compared to those who ignored the prime word. With respect to our post-hoc explanations (attention alignment for the ambiguity dimension and possibly classifying the prime regarding that dimension which could lead to a segmentation of the mental stream of information) for the failed replication of learning which words are ambiguous and which words have only one meaning prior to the EP paradigm (study 3), a further experimental investigation of the influence of attention in general (how much attention is aligned to the prime) for the multiple categorizable primes (ambiguity and valence dimension) and in particular the influence of directing attention to the prime feature of ambiguity would be interesting and might contribute to a better understanding of our results.

Moreover, the role of consciousness for the ambiguousness of words would be interesting: Does the mental state of having one vs. two meanings in mind needs to be consciously processed in order to trigger an evaluation or does it inhibit an evaluative response? Regarding the results form Gawronski et al. (2010) which showed that the EP paradigm but not the AMP (Payne et al., 2005) was sensitive to manipulations of attention to specific prime features, this indirect measure (although less reliable than the EP paradigm) could also be an attractive option in order to investigate unintentional activation of evaluation of ambiguity.

Based on Moors et al. (2010) notion that two sequential processes take place in indirect tasks, that is the evaluation and the translation of the evaluation in direct observable responses, Deutsch and Gawronski (2009) provided evidence that the same experimental manipulation (by differently affecting the translation process) can produce different effects on two different priming based indirect measures. In that sense, a manipulation aiming to change the attitude towards an object can be misattributed as being effective although only task-specific mechanisms relevant to translate the (unaffected) evaluation are in fact manipulated. In the same way a superficially ineffective attitude change manipulation (as indicated by the observed responses) could be indeed effective (in a sense that the evaluations changed) but masked by likewise changed translation processes. In that sense, our attempts to experimentally manipulate the automatic evaluations of ambiguity and clearness via the IAT in studies 5 and 6 also introduce an additional source of uncertainty when interpreting the results obtained in the EP paradigm. As our studies designed to manipulate the attitude towards ambiguity prior to the EP paradigm, revealed contrastive (study 5) or no effects (study 6) in the EP paradigm, a more subtle manipulation of attitude towards ambiguity and clearness prior to the EP paradigm would potentially be better suited. Specifically, a manipulation without potential associations between concepts and response sides could be more suitable, as this probably decreases the influence of the attitude manipulation on the translation phase of the EP paradigm, for which there is evidence that it is mainly driven by response interference processes (Klauer, 1997; Klauer et al., 2005; Klinger et al., 2000; Voss et al., 2013). For instance, evaluative learning via evaluative conditioning could be an attractive alternative in order to change the evaluations of subjective ambiguity, although it is still discussed which processes (e.g. associative or propositional) are involved in evaluative conditioning (Hütter & Rothermund, 2020).

One could argue that the state of having two meanings of a lexically ambiguous word in mind is like the simultaneous processing of two lexically unambiguous words, which also leads to two distinct meanings in mind. In this respect one could also investigate if the simultaneous activation of two meanings of two lexically unambiguous words trigger a different evaluation than the single activation of one meaning from one word (or two same words). If that would be the case, ambiguous stimuli would only be a parsimonious way to achieve a state of simultaneous activation of several meanings/interpretations or generally speaking representations. However, we can assume that the state of having multiple distinct representations in mind derived from a single stimulus is different from the one derived from several stimuli in a sense that the former indirectly implies (in most contexts) that only one representation is correct which may result in a process to identify the correct one. Our paradigm, however, provided no information which could be used for disambiguation. The reason for that was that we were interested in evaluation of not-disambiguated ambiguity. One could argue that, as there is no cue for disambiguation, there is no need for disambiguation. Further investigations should focus on potential differences in evaluation of ambiguity with and without the possibility/necessity to disambiguate. We already mentioned that being ambivalent towards an object only triggers discomfort when we have to make a decision about the object (Nohlen et al., 2016; van Harreveld et al., 2009). In a similar vein, it could be that the need to select one meaning is a moderator of the strength (or even direction) of the evaluation of the state of having multiple meanings in mind. This could be investigated by either manipulating or measuring the need to select the "correct" meaning.

Regarding the mentioned role of potential fluency differences between ambiguous and unambiguous words, attempting to rule this out in future studies would also contribute to a clearer interpretation of results in terms of indicating different automatic activations of evaluation between ambiguous and unambiguous words only based on a difference in ambiguity and not in processing fluency. A learning phase, in which the meanings of ambiguous and unambiguous words are pre-activated could potentially reduce differences in processing fluency. However, we already argued (see section 3.5.4) that this may also increase the attention to the discriminative feature of ambiguity of the primes in the EP paradigm. Moreover, learning which word is ambiguous and which word only has one meaning prior to the EP paradigm possibly establishes associations of the specific word with the labels ambiguity and clearness which prevents the interpretation of potential EP effects in terms of reflecting different automatic activations of evaluations between the state of having one vs. two meanings in mind.

Finally, De Houwer et al. (2020) mentioned the influence of (missing) relational information for indirect tasks. To avoid that participants apply different relations (see 4.1) between attributes (e.g. positive and negative) and concepts (e. g. ambiguity and clearness), it could be helpful to include relational information by using terms in the IAT like "I like/I dislike" instead of "positive/negative" in order to reduce unintended variance.

4.4 Relevance

Although there is evidence that *what* is assessed with indirect measures is not based on a different learning basis than what is assessed with explicit measures, as these types of measures are not only sensitive to one specific learning history (Corneille & Hütter, 2020), there are still good reasons to assess the evaluative response towards subjective ambiguity by indirect measures. De Houwer (2008) argue that the predictive value of indirect measures for real-life behavior depends on the overlap of functional properties (conditions under which the evaluative responding took place – similar to the term *operating conditions* used by (Gawronski & Bodenhausen, 2009, 2014)). As subjective ambiguity is a short-time phenomenon (the context in real-life usually quickly resolves the ambiguity which also often occurs unintentionally in real-life) the evaluative responses to subjective ambiguity in real-life are probably more similar to the evaluative responses in indirect (response time based) tasks (especially the EP paradigm) compared to direct measures. Because of this assumed overlap of functional properties, indirect measures provide better insights into the evaluative processing of subjective ambiguity. For an extensive overview of conditions when indirect and when direct measures have predictive validity for certain types of behavior, see Perugini, Richetin, and Zogmaister (2010).

But why should it be important at all to know the attitude towards ambiguity in general or towards subjective ambiguity in particular? Eagly and Chaiken (1993; p.1) pointed out that "the discrepant attitudes that often characterize different subgroups of a society are believed to underlie the social conflict that political and social issues sometimes engender". Assuming that ambiguity is a key element of complexity in modern life, different evaluative reactions to ambiguity could even play a role in how a society becomes polarized. Since Adorno et al. (1950) provided evidence that politically right-wing oriented persons are less tolerant of ambiguity, a lot of work has been done on the relationship between political or sociocultural attitudes and TA or related attitudes. In a meta-analysis, Van Hiel, Onraet, and De Pauw (2010) reported a mean r of .22 (based on 20 studies) between intolerance of ambiguity and indicators of right-wing political orientation. Of course, these studies investigated correlations with the concept of TA which has a broader conceptualization than our defined attitude towards ambiguity. However, the assessment of unintentional and fast evaluative responses to ambiguity might also have predictive power for more spontaneous, impulsive real-life behaviors that can be located on the political spectrum. For instance, gender-ambiguous and ethnic-ambiguous people potentially trigger different spontaneous real-life evaluative responses corresponding to the assessed (via indirect measures) evaluative responses to subjective ambiguity. In that sense, being aware of one's own individual attitude towards ambiguity might offer an opportunity to deliberately control for the (e.g. social) consequences of these automatic evaluations.

If we consider the mental state of having multiple distinct representations in mind derived from a single stimulus as a specific form of representation which is different (and which potentially triggers a different evaluation) from the representation constituted by the fact of having multiple distinct representations in mind derived from multiple stimuli, it could be worthwhile to think of the nature of this superordinate representation. In attitude research, two forms of representations guiding the evaluation are discussed: associational representations and propositional/symbolic representations (for comparison of different forms of representations see Hummel, 2010). Propositional representations differ from simple associations as they encode information about how elements are related (e.g. X is *bigger than* Y) and have a truth value (De Houwer et al., 2020). The superordinate representation of having multiple distinct representations (A *and* B) in mind derived from a single stimulus is possibly propositional as one could argue that it has a truth value (which is false) as only one representation (A *or* B) can be true at a certain point in time in most contexts. However, considerations of the nature of representation of the attitude object (subjective ambiguity) itself does not imply whether the attitude towards this object is associative (e.g., Fazio, 1990, 2001; Fazio, 2007; Fazio et al., 1986) or propositional (e.g., De Houwer et al., 2020) or both kinds (e.g., R. J. Rydell & McConnell, 2006 ; but see Heycke, Gehrmann, Haaf, & Stahl (2018) for a failed replication) represented. What kind of representations mediate the evaluation of this possibly propositionally represented attitude object goes beyond the scope of the present work.

Assuming that the state of having multiple distinct representations in mind derived from a single stimulus is an instance of propositional and amodal representation, subjective ambiguity can be considered an abstract representation according to the first sense (modal = concrete, amodal = abstract) of abstractness reported in the taxonomy of abstraction by Reed (2016). For concrete (modally represented) attitude objects it can be assumed that the activated patterns of representations derived from verbal descriptions of the object and those derived from perceiving the attitude object might be more similar. For instance, it can be assumed that the evaluative response to the question “do you like chocolate” and the evaluative response to a piece of chocolate are similar. For attitude objects for which there is no single physically existing referent, and which can be regarded as abstract in that sense, like “religion” or “democracy”, there is obviously no single object (probably except for the words referring to the concept) that have the potential to activate the complete representation of the attitude object (as they are qualitatively and quantitatively different for each individual). However, subjective ambiguity is an exceptionally specific attitude object (in that it is the same for each person) characterized by the mental state of having multiple distinct interpretations in mind. This highly specific attitude object can be activated by any single ambiguous (and not disambiguated) item. In that sense, one could argue that subjective ambiguity is neither a concrete attitude object (e.g., chocolate) having a material referent and being represented modally, nor does it resemble other abstract attitude objects (e.g., religion) for which no single information (e.g., the Cross) has the potential to activate the representation completely. This special status of subjective ambiguity together with the found evidence for evaluative responding to it highlights the assumed universal occurrence of evaluation.

4.5 Conclusion

The unsatisfactory validation approaches of direct measures of (in)tolerance of ambiguity (unsatisfactory as it was never demonstrated that direct measures predict evaluative responses at the moment of being in an ambiguous situation) was the starting point of this thesis and inspired us to investigate via indirect measures whether there is evidence of evaluative responding when having multiple interpretations in mind at all, and in particular, constituting our main question, whether subjective ambiguity (which we defined as the mental state of having multiple meanings in mind) could trigger, unintentionally and fast, interindividually different evaluations, thus indicating substantial variation in liking. Although, admittedly, the evidence for interindividual differences in automatic activation of evaluations of the mental state characterized by the activation of two (vs. one) meanings of lexical ambiguous words is weak, results of the EP paradigm provided evidence for unintentional fast activation of evaluations differing in terms of valence for lexical ambiguous compared to lexical unambiguous words. In general, the results of the EP paradigm speak for a more negative (or less positive) unintentional and fast evaluation of lexical ambiguous compared to lexical unambiguous words. As these two types of stimuli were selected to differ only in the number of meanings (one vs. two), it can be assumed that the mental state of having two meanings in mind (vs. only one meaning in mind) triggers a more negative (or less positive) evaluation.

Furthermore, the IAT results of all studies (except for study 2, in which the IAT *D* score was close to zero) provide evidence that, on average, the associations of ambiguity with negative valence and clearness with positive valence are stronger than the associations of ambiguity with positive valence and clearness with negative valence. Assuming that the activation of one meaning and thus clearness (only in the specific sense of being the opposite of ambiguity) is the default state, we could further hold that for these entities an evaluative response (and an association with valence) is uneconomic and hence unlikely. Thus, our data could be interpreted to mean that generally (for our investigated samples): (1) the mental state of having multiple meanings in mind trigged unintentional and fast a negative evaluation and (2) ambiguity is associated with negativity.

Coming back to the question raised at the beginning of this thesis whether there is a specific attitude determining our evaluative responses towards ambiguous stimuli, we can answer with reference to the definition of attitude provided by Eagly and Chaiken (1993) that we found “a general psychological tendency that is expressed by evaluating a particular entity [in our case: subjective ambiguity] with some degree of […] disfavor” (p. 1). Although, we found only limited evidence for interindividual different unintentional and fast activations of evaluations of ambiguity, we can say that there is evidence for a negative attitude towards subjective ambiguity at a group level. However, whether there is a general attitude towards ambiguity which determines the evaluative responses towards several types of ambiguity (e.g., ambiguity at the level of sentences like garden path, ambiguity on the level of passages, ambiguity within social situations etc.) cannot be answered with our data from indirect measures, as we only investigated the automatic evaluative responses towards lexically ambiguous vs. unambiguous words. While this approach fits our research question, it is obviously not suited to investigate the evaluative responses to other types of ambiguity which need more deliberate processing in order to carve out the meanings.

Being aware of the limited relevance of our results from the indirect measures with respect to ambiguity at other levels or in other domains, we investigated the structure of attitude towards ambiguity as well as developed a new direct measure. Thereby we found evidence for a predominantly domain insensitive attitude towards ambiguity, except for the domain art. The factors capturing the attitude towards ambiguity in general and the attitude towards ambiguity in art showed a moderate correlation. We discussed this in terms of centrality of ambiguity in art vs. other domains. In art, ambiguity is a more central element and the confrontation with this kind of ambiguity can be expected when perceiving art. Although we employed several approaches to validate these direct measures of attitude towards ambiguity by showing that they predict evaluative responses in an ambiguous situation, additional work on these validation attempts is needed in order to answer the question of whether there is a general evaluative tendency towards ambiguity.

By providing evidence that evaluative responding to the mental state of having multiple meanings in mind occurs, this thesis can be understood as a starting point to the investigation of evaluation of the mental representation of ambiguity.

References

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. (1950). *The Authoritarian personality*. New York: Harper.

Albarracín, D., Hart, W., & McCulloch, K. C. (2006). Associating versus proposing or associating what we propose: Comment on Gawronski and Bodenhausen (2006). *Psychological Bulletin, 132*(5), 732-735. doi:10.1037/0033-2909.132.5.732

Alexopoulos, T., Fiedler, K., & Freytag, P. (2012). The impact of open and closed mindsets on evaluative priming. *Cognition and Emotion, 26*(6), 978-994. doi:10.1080/02699931.2011.630991

Alexopoulos, T., Lemonnier, A., & Fiedler, K. (2017). Higher order influences on evaluative priming: Processing styles moderate congruity effects. *Cognition and Emotion, 31*(1), 57-68. doi:10.1080/02699931.2015.1080666

Allen, T. J., Sherman, J. W., & Klauer, K. C. (2010). Social context and the self-regulation of implicit bias. *Group Processes & Intergroup Relations, 13*(2), 137-149. doi:10.1177/1368430209353635

Allport, G. W. (1935). Attitudes. In C. Murchinson (Ed.), *A Handbook of Social Psychology*. (pp. 798-844). Worcester, MA: Clark University Press.

Altarriba, J., & Gianico, J. L. (2003). Lexical ambiguity resolution across languages: A theorical and empirical review. *Experimental Psychology, 50*(3), 159-170. doi:10.1026//1617-3169.50.3.159

Augustinova, M., & Ferrand, L. (2014). Automaticity of Word Reading: Evidence From the Semantic Stroop Paradigm. *Current Directions in Psychological Science, 23*(5), 343-348. doi:10.1177/0963721414540169

Back, M. D., Schmukle, S. C., & Egloff, B. (2005). Measuring Task-Switching Ability in the Implicit Association Test. *Experimental Psychology, 52*(3), 167-179. doi:10.1027/1618-3169.52.3.167

Bading, K., Stahl, C., & Rothermund, K. (2020). Why a standard IAT effect cannot provide evidence for association formation: the role of similarity construction. *Cognition and Emotion, 34*(1), 128-143. doi:10.1080/02699931.2019.1604322

Bagozzi, R. P. (1978). The Construct Validity Of The Affective, Behavioral, And Cognitive Components Of Attitude By Analysis Of Covariance Structures. *Multivariate Behavioral Research, 13*(1), 9-31. doi:10.1207/s15327906mbr1301_2

Balcetis, E., & Dunning, D. (2006). See what you want to see: Motivational influences on visual perception. *Journal of Personality and Social Psychology, 91*(4), 612-625. doi:10.1037/0022-3514.91.4.612

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual Word Recognition of Single-Syllable Words. *Journal of Experimental Psychology: General, 133*(2), 283-316. doi:10.1037/0096-3445.133.2.283

Banse, R., Seise, J., & Zerbes, N. (2001). Implicit Attitudes towards Homosexuality: Reliability, Validity, and Controllability of the IAT. *Experimental Psychology, 48*(2), 145-160. doi:10.1026//0949-3946.48.2.145

Bar-Anan, Y., & Nosek, B. A. (2014). A comparative investigation of seven indirect attitude measures. *Behavior Research Methods, 46*(3), 668-688. doi:10.3758/s13428-013-0410-6

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition: Basic processes; Applications., Vols. 1-2, 2nd ed.* (pp. 1-40). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology, 62*(6), 893-912. doi:10.1037/0022-3514.62.6.893

Bargh, J. A., & Pietromonaco, P. (1982). Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology, 43*(3), 437-449. doi:10.1037/0022-3514.43.3.437

Bartlett, M. S. (1950). TESTS OF SIGNIFICANCE IN FACTOR ANALYSIS. *British Journal of Statistical Psychology, 3*(2), 77-85. doi:10.1111/j.2044-8317.1950.tb00285.x

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1).

Bauer, M., Knape, J., Koch, P., & Winkler, S. (2010). Dimensionen der Ambiguität. *Zeitschrift für Literaturwissenschaft und Linguistik, 158*, 7 -75.

Benjamin Jr, A. J., Riggio, R. E., & Mayes, B. T. (1996). Reliability and factor structure of Budner's Tolerance for ambiguity scale. *Journal of Social Behavior & Personality, 11*(3), 625-632.

Beretta, A., Fiorentino, R., & Poeppel, D. (2005). The effects of homonymy and polysemy on lexical access: an MEG study. *Cognitive Brain Research, 24*(1), 57-65. doi:http://dx.doi.org/10.1016/j.cogbrainres.2004.12.006

Berger, K. T., Hütter, M., & Corneille, O. (2019). Investigating attitudinal ambivalence via sequential priming: Evidence for the simultaneous and unintentional activation of opposite evaluations. *Journal of Experimental Psychology: General*. doi:10.1037/xge0000572

Bilenko, N. Y., Grindrod, C. M., Myers, E. B., & Blumstein, S. E. (2009). Neural correlates of semantic competition during processing of ambiguous words. *Journal of Cognitive Neuroscience, 21*(5), 960-975. doi:10.1162/jocn.2009.21073

Black, M. (1937). Vagueness. An Exercise in Logical Analysis. *Philosophy of Science, 4*(4), 427-455.

Blair, I. V. (2002). The Malleability of Automatic Stereotypes and Prejudice. *Personality and Social Psychology Review, 6*(3), 242-261. doi:10.1207/S15327957PSPR0603_8

Bless, H., Wänke, M., Bohner, G., & Fellhauer, R. F. (1994). Need for Cognition: Eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben = Need for cognition: A scale measuring engagement and happiness in cognitive tasks. *Zeitschrift für Sozialpsychologie, 25*(2), 147-154.

Bohner, G., Chaiken, S., & Hunyadi, P. (1994). The role of mood and message ambiguity in the interplay of heuristic and systematic processing. *European Journal of Social Psychology, 24*(1), 207-221. doi:10.1002/ejsp.2420240115

Bohner, G., & Dickel, N. (2011). Attitudes and Attitude Change. *Annual Review of Psychology, 62*(1), 391-417. doi:10.1146/annurev.psych.121208.131609

Bors, D. A., Gruman, J. A., & Shukla, S. (2010). Measuring tolerance of ambiguity: Item polarity, dimensionality, and criterion validity. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology, 60*(4), 239-245. doi:http://dx.doi.org/10.1016/j.erap.2010.07.001

Bosson, J. K., Swann, W. B., Jr., & 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*(4), 631-643. doi:10.1037/0022-3514.79.4.631

Bradley, M., & Lang, P. J. (1999). *The International Affective Digitized Sounds (IADS): Stimuli, instruction manual and affective ratings. Technical Report No. C-1.* Center for research in psychophysiology, University of Florida, Gainesville, FL, Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49-59. doi:http://dx.doi.org/10.1016/0005-7916(94)90063-9

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings. Technical report C-1*. Center for research in psychophysiology, University of Florida, Gainesville, FL.

Breckler, S. J. (1984). Empirical validation of affect, behavior, and cognition as distinct components of attitude. *Journal of Personality and Social Psychology, 47*(6), 1191-1205. doi:10.1037/0022-3514.47.6.1191

Breckler, S. J., & Wiggins, E. C. (1989). Affect versus evaluation in the structure of attitudes. *Journal of Experimental Social Psychology, 25*(3), 253-271. doi:https://doi.org/10.1016/0022-1031(89)90022-X

Briesemeister, B. B., Kuchinke, L., & Jacobs, A. M. (2011). Discrete emotion norms for nouns: Berlin affective word list (DENN–BAWL). *Behavior Research Methods, 43*(2), 441-448. doi:10.3758/s13428-011-0059-y

Briñol, P., Petty, R. E., & McCaslin, M. J. (2008). Changing attitudes on implicit versus explicit measures: What is the difference? In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures.* (pp. 285-326). New York, NY: Psychology Press.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology, 58*(5), 412-424. doi:10.1027/1618-3169/a000123

Budner, S. (1962). Intolerance of ambiguity as a personality variable. *Journal of Personality, 30*(1), 29. doi:10.1111/1467-6494.ep8933446

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin, 115*(3), 401-423. doi:10.1037/0033-2909.115.3.401

Caglioti, G. (1992). *The Dynamics of Ambiguity*. Berlin: Springer.

Calanchini, J., Sherman, J. W., Klauer, K. C., & Lai, C. K. (2014). Attitudinal and Non-Attitudinal Components of IAT Performance. *Personality and Social Psychology Bulletin, 40*(10), 1285-1296. doi:10.1177/0146167214540723

Caligiuri, P. M., Jacobs, R. R., & Farr, J. L. (2000). The Attitudinal and Behavioral Openness Scale: scale development and construct validation. *International Journal of Intercultural Relations, 24*(1), 27-46. doi:https://doi.org/10.1016/S0147-1767(99)00021-8

Cameron, C. D., Brown-Iannuzzi, J. L., & Payne, B. K. (2012). Sequential Priming Measures of Implicit Social Cognition. *Personality and Social Psychology Review, 16*(4), 330-350. doi:10.1177/1088868312440047

Chaiken, S., & Maheswaran, D. (1994). Heuristic Processing Can Bias Systematic Processing: Effects of Source Credibility, Argument Ambiguity, and Task Importance on Attitude Judgment. *Journal of Personality & Social Psychology, 66*(3), 460-473.

Chan, E., Ybarra, O., & Schwarz, N. (2006). Reversing the affective congruency effect: The role of target word frequency of occurrence. *Journal of Experimental Social Psychology, 42*(3), 365-372. doi:https://doi.org/10.1016/j.jesp.2005.04.008

Chen, M., & Bargh, J. A. (1999). Consequences of Automatic Evaluation: Immediate Behavioral Predispositions to Approach or Avoid the Stimulus. *Personality and Social Psychology Bulletin, 25*(2), 215-224. doi:10.1177/0146167299025002007

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

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204-256. doi:10.1037/0033-295X.108.1.204

Cone, J., Mann, T. C., & Ferguson, M. J. (2017). Chapter Three - Changing Our Implicit Minds: How, When, and Why Implicit Evaluations Can Be Rapidly Revised. In J. M. Olson (Ed.), *Advances in Experimental Social Psychology* (Vol. 56, pp. 131-199): Academic Press.

Conner, M., & Sparks, P. (2002). Ambivalence and Attitudes. *European Review of Social Psychology, 12*(1), 37-70. doi:10.1080/14792772143000012

Conrad, C. (1974). Context effects in sentence comprehension: A study of the subjective lexicon. *Memory & Cognition, 2*(1), 130-138. doi:10.3758/BF03197504

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating Multiple Processes in Implicit Social Cognition: The Quad Model of Implicit Task Performance. *Journal of Personality and Social Psychology, 89*(4), 469-487. doi:10.1037/0022-3514.89.4.469

Corneille, O., & Hütter, M. (2020). Implicit? What Do You Mean? A Comprehensive Review of the Delusive Implicitness Construct in Attitude Research. *Personality and Social Psychology Review*, *24*(3), 212-232. doi:10.1177/1088868320911325

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology, 78*(1), 98.

Craeynest, M., Crombez, G., Koster, E. H. W., Haerens, L., & De Bourdeaudhuij, I. (2008). Cognitive-motivational determinants of fat food consumption in overweight and obese youngsters: The implicit association between fat food and arousal. *Journal of Behavior Therapy and Experimental Psychiatry, 39*(3), 354-368. doi:https://doi.org/10.1016/j.jbtep.2007.09.002

Crocker, J., Cornwell, B., & Major, B. (1993). The stigma of overweight: Affective consequences of attributional ambiguity. *Journal of Personality and Social Psychology, 64*(1), 60-70. doi:10.1037/0022-3514.64.1.60

Crocker, J., Voelkl, K., Testa, M., & Major, B. (1991). Social stigma: The affective consequences of attributional ambiguity. *Journal of Personality and Social Psychology, 60*(2), 218-228. doi:10.1037/0022-3514.60.2.218

Croizet, J.-C., & Fiske, S. T. (2000). Moderation of Priming by Goals: Feeling Entitled to Judge Increases Judged Usability of Evaluative Primes. *Journal of Experimental Social Psychology, 36*(2), 155-181. doi:https://doi.org/10.1006/jesp.1999.1397

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: consistency, stability, and convergent validity. *Psychological Science, 12*(2), 163-170.

Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations: a social cognitive neuroscience perspective. *Trends in Cognitive Sciences, 11*(3), 97-104. doi:https://doi.org/10.1016/j.tics.2006.12.005

Dalbert, C. (1999). Die Ungewißheitstoleranzskala: Skaleneigenschaften und Validierungsbefunde. *Hallesche Berichte zur Pädagogischen Psychologie, Nr. 1*.

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*(5), 800-814. doi:10.1037/0022-3514.81.5.800

De Houwer, J. (2001). A Structural and Process Analysis of the Implicit Association Test. *Journal of Experimental Social Psychology, 37*(6), 443-451. doi:https://doi.org/10.1006/jesp.2000.1464

De Houwer, J. (2003a). The extrinsic affective Simon task. *Experimental Psychology, 50*(2), 77-85. doi:10.1026//1618-3169.50.2.77

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 Associates Publishers.

De Houwer, J. (2008). Comparing measures of attitudes at the functional and procedural level: Analysis and implications. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures.* (pp. 361-390). New York: Psychology Press.

De Houwer, J. (2009). How do People Evaluate Objects? A Brief Review. *Social and Personality Psychology Compass*, *3*(1), 36-48. doi:10.1111/j.1751-9004.2008.00162.x

De Houwer, J., Crombez, G., Koster, E. H. W., & Beul, N. D. (2004). Implicit alcohol-related cognitions in a clinical sample of heavy drinkers. *Journal of Behavior Therapy and Experimental Psychiatry, 35*(4), 275-286. doi:https://doi.org/10.1016/j.jbtep.2004.05.001

De Houwer, J., & De Bruycker, E. (2007). The implicit association test outperforms the extrinsic affective Simon task as an implicit measure of inter-individual differences in attitudes. *British Journal of Social Psychology, 46*(2), 401.

De Houwer, J., & Eelen, P. (1998). An Affective Variant of the Simon Paradigm. *Cognition and Emotion, 12*(1), 45-62. doi:10.1080/026999398379772

De Houwer, J., Geldof, T., & De Bruycker, E. (2005). The Implicit Association Test as a General Measure of Similarity. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 59*(4), 228-239. doi:10.1037/h0087478

De Houwer, J., Hermans, D., Rothermund, K., & Wentura, D. (2002). Affective priming of semantic categorisation responses. *Cognition and Emotion, 16*(5), 643-666. doi:10.1080/02699930143000419

De Houwer, J., Van Dessel, P., & Moran, T. (2020). Chapter Three - Attitudes beyond associations: On the role of propositional representations in stimulus evaluation. In B. Gawronski (Ed.), *Advances in Experimental Social Psychology* (Vol. 61, pp. 127-183): Academic Press.

de la Vega, I., Dudschig, C., De Filippis, M., Lachmair, M., & Kaup, B. (2013). Keep your hands crossed: The valence-by-left/right interaction is related to hand, not side, in an incongruent hand–response key assignment. *Acta Psychologica, 142*(2), 273-277. doi:https://doi.org/10.1016/j.actpsy.2012.12.011

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods, 47*(1), 1-12. doi:10.3758/s13428-014-0458-y

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*(3), 371-396. doi:10.1037/0096-3445.122.3.371

Deutsch, R., & Gawronski, B. (2009). When the method makes a difference: Antagonistic effects on “automatic evaluations” as a function of task characteristics of the measure. *Journal of Experimental Social Psychology, 45*(1), 101-114. doi:https://doi.org/10.1016/j.jesp.2008.09.001

Diener, E., Larsen, R. J., & Emmons, R. A. (1984). Person×Situation interactions: Choice of situations and congruence response models. *Journal of Personality and Social Psychology, 47*(3), 580-592. doi:10.1037/0022-3514.47.3.580

Dijksterhuis, A., & Aarts, H. (2003). On Wildebeests and Humans: The Preferential Detection of Negative Stimuli. *Psychological Science, 14*(1), 14-18. doi:10.1111/1467-9280.t01-1-01412

Duckworth, K. L., Bargh, J. A., Garcia, M., & Chaiken, S. (2002). The automatic evaluation of novel stimuli. *Psychological Science, 13*(6), 513-519. doi:10.1111/1467-9280.00490

Duffy, S. A., Kambe, G., & Rayner, K. (2001). The effect of prior disambiguating context on the comprehension of ambiguous words: Evidence from eye movements. In D. S.

Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity*. (pp. 27-43). Washington, DC: American Psychological Association.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27*(4), 429-446. doi:http://dx.doi.org/10.1016/0749-596X(88)90066-6

Duffy, S. E., & Feist, M. I. (2014). Individual differences in the interpretation of ambiguous statements about time. *Cognitive Linguistics, 25*(1), 29-54. doi:10.1515/cog-2013-0030

Durrheim, K., & Foster, D. (1997). Tolerance of ambiguity as a content specific construct. *Personality and Individual Differences, 22*(5), 741-750. doi:http://dx.doi.org/10.1016/S0191-8869(96)00207-3

Dziuban, C. D., & Shirkey, E. C. (1974). When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychological Bulletin, 81*(6), 358.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Orlando, FL: Harcourt Brace Jovanovich College Publishers.

Eagly, A. H., & Chaiken, S. (2007). The advantages of an inclusive definition of attitude. *Social Cognition, 25*(5), 582-602. doi:10.1521/soco.2007.25.5.582

Ebeling, K. S., & Spear, P. S. (1980). Preference and performance on two tasks of varying ambiguity as a function of ambiguity tolerance. *Australian Journal of Psychology, 32*(2), 127-133. doi:10.1080/00049538008254679

Ebert, I. D., Steffens, M. C., Stülpnagel, R. v., & Jelenec, P. (2009). How to like yourself better, or chocolate less: Changing implicit attitudes with one IAT task. *Journal of Experimental Social Psychology, 45*(5), 1098-1104. doi:https://doi.org/10.1016/j.jesp.2009.06.008

Echterhoff, G., Higgins, E. T., & Levine, J. M. (2009). Shared Reality: Experiencing Commonality with others' Inner States about the World. *Perspectives on Psychological Science, 4*(5), 496-521. doi:10.1111/j.1745-6924.2009.01161.x

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an implicit association test for assessing anxiety. *Journal of Personality and Social Psychology, 83*(6), 1441-1455. doi:10.1037/0022-3514.83.6.1441

Ehrlich, D. (1965). “Intolerance of Ambiguity,” Walk's a Scale: Historical Comment. *Psychological Reports, 17*(2), 591-594. doi:10.2466/pr0.1965.17.2.591

Ellsberg, D. (1961). RISK, AMBIGUITY, AND THE SAVAGE AXIOMS. *Quarterly Journal of Economics, 75*(4), 643-669. doi:10.2307/1884324

Ely, C. M. (1989). Tolerance of ambiguity and use of second language strategies. *Foreign Language Annals, 22*(5), 437 - 446.

Endres, M. L., Camp, R., & Milner, M. (2015). Is ambiguity tolerance malleable? Experimental evidence with potential implications for future research. *Frontiers in Psychology, 6*(619). doi:10.3389/fpsyg.2015.00619

English, H. B., & English, A. C. (1958). *A comprehensive dictionary of psychological and psychoanalytical terms: A guide to usage*. Oxford: Longmans, Green.

Erdley, C. A., & D'Agostino, P. R. (1988). Cognitive and affective components of automatic priming effects. *Journal of Personality and Social Psychology, 54*(5), 741-747. doi:10.1037/0022-3514.54.5.741

Eysenck, M. W., MacLeod, C., & Mathews, A. (1987). Cognitive functioning and anxiety. *Psychological Research, 49*(2), 189-195. doi:10.1007/BF00308686

Eysenck, M. W., Mogg, K., May, J., Richards, A., & Mathews, A. (1991). Bias in interpretation of ambiguous sentences related to threat in anxiety. *Journal of Abnormal Psychology, 100*(2), 144-150. doi:10.1037/0021-843X.100.2.144

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): Academic Press.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion*, *15*(2), 115-141. doi:10.1080/02699930125908

Fazio, R. H. (2007). Attitudes as Object–Evaluation Associations of Varying Strength. *Social Cognition, 25*(5), 603-637. doi:10.1521/soco.2007.25.5.603

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69*(6), 1013-1027. doi:10.1037/0022-3514.69.6.1013

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*(2), 229-238. doi:10.1037/0022-3514.50.2.229

Ferguson, M. J., Bargh, J. A., & Nayak, D. A. (2005). After-affects: How automatic evaluations influence the interpretation of subsequent, unrelated stimuli. *Journal of Experimental Social Psychology, 41*(2), 182-191. doi:10.1016/j.jesp.2004.05.008

Fiedler, K., Bluemke, M., & Unkelbach, C. (2011). On the adaptive flexibility of evaluative priming. *Memory & Cognition, 39*(4), 557-572. doi:10.3758/s13421-010-0056-x

Fiedler, K., & Hütter, M. (2014). The limits of automaticity. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind.* (pp. 497-513). New York, NY: The Guilford Press.

Forster, K. I., & Bednall, E. S. (1976). Terminating and exhaustive search in lexical access. *Memory & Cognition*, *4*(1), 53-61. doi:10.3758/BF03213255

Frenck-Mestre, C., & Prince, P. (1997). Second Language Autonomy. *Journal of Memory and Language*, *37*(4), 481-501. doi:https://doi.org/10.1006/jmla.1997.2526

Frenkel-Brunswik, E. (1948). Tolerance toward ambiguity as a personality variable. *American Psychologist, 3*, 268.

Frenkel-Brunswik, E. (1949). Intolerance of ambiguity as an emotional perceptual personality variable. *Journal of Personality, 18*(1), 108-143. doi:10.1111/j.1467-6494.1949.tb01236.x

Frings, C., & Wentura, D. (2008). Trial-by-trial effects in the affective priming paradigm. *Acta Psychologica*, *128*(2), 318-323. doi:https://doi.org/10.1016/j.actpsy.2008.03.004

Furnham, A. (1994). A content, correlational and factor analytic study of four tolerance of ambiguity questionnaires. *Personality and Individual Differences, 16*(3), 403-410. doi:https://doi.org/10.1016/0191-8869(94)90066-3

Furnham, A., & Avison, M. (1997). Personality and preference for surreal paintings. *Personality and Individual Differences, 23*(6), 923-935. doi:https://doi.org/10.1016/S0191-8869(97)00131-1

Furnham, A., & Marks, J. (2013). Tolerance of Ambiguity: A Review of the Recent Literature. *Psychology, Vol.04No.09*, 12. doi:10.4236/psych.2013.49102

Furnham, A., & Ribchester, T. (1995). Tolerance of ambiguity: A review of the concept, its measurement and applications. *Current Psychology, 14*(3), 179-199. doi:10.1007/bf02686907

Gawlick-Grendell, L. A., & Woltz, D. J. (1994). Meaning dominance norms for 120 homographs. *Behavior Research Methods, Instruments, & Computers, 26*(1), 5-25. doi:10.3758/BF03204557

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*(5), 692-731. doi:10.1037/0033-2909.132.5.692

Gawronski, B., & Bodenhausen, G. V. (2009). Operating principles versus operating conditions in the distinction between associative and propositional processes. *Behavioral and Brain Sciences*, *32*(2), 207-208. doi:10.1017/S0140525X09000958

Gawronski, B., & Bodenhausen, G. V. (2011). Chapter two - The Associative–Propositional Evaluation Model: Theory, Evidence, and Open Questions. In J. M. Olson & M. P. Zanna (Eds.), *Advances in Experimental Social Psychology* (Vol. 44, pp. 59-127): Academic Press.

Gawronski, B., & Bodenhausen, G. V. (2014). The associative—Propositional evaluation model: Operating principles and operating conditions of evaluation. In J. W. Sherman, B. Gawronski, Y. Trope, J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind.* (pp. 188-203). New York, NY, US: The Guilford Press.

Gawronski, B., Cunningham, W. A., LeBel, E. P., & Deutsch, R. (2010). Attentional influences on affective priming: Does categorisation influence spontaneous evaluations of multiply categorisable objects? *Cognition and Emotion, 24*(6), 1008-1025. doi:10.1080/02699930903112712

Geller, G., Tambor, E. S., Chase, G. A., & Holtzman, N. A. (1993). Measuring Physicians' Tolerance for Ambiguity and Its Relationship to Their Reported Practices regarding Genetic Testing. *Medical Care, 31*(11), 989-1001.

Gibbons, H. (2009). Evaluative priming from subliminal emotional words: Insights from event-related potentials and individual differences related to anxiety. *Consciousness and Cognition, 18*(2), 383-400. doi:https://doi.org/10.1016/j.concog.2009.02.007

Glaser, J. (2003). Reverse priming: Implications for the (un)conditionality of automatic evaluation. In *The psychology of evaluation: Affective processes in cognition and emotion.* (pp. 87-108). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Glover, J. A., Romero, D., Romero, P., & Petersen, C. (1978). EFFECTS OF A SIMULATION GAME UPON TOLERANCE FOR AMBIGUITY, DOGMATISM, AND RISK TAKING. *Journal of Social Psychology, 105*(2), 291. doi:10.1080/00224545.1978.9924126

Glucksberg, S., Kreuz, R. J., & Rho, S. H. (1986). Context can constrain lexical access: Implications for models of language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*(3), 323-335. doi:10.1037/0278-7393.12.3.323

Govan, C. L., & Williams, K. D. (2004). Changing the affective valence of the stimulus items influences the IAT by re-defining the category labels. *Journal of Experimental Social Psychology, 40*(3), 357-365. doi:https://doi.org/10.1016/j.jesp.2003.07.002

Graf, R., Nagler, M., & Jacobs, A. M. (2005). Faktorenanalyse von 57 Variablen der visuellen Worterkennung. *Zeitschrift für Psychologie / Journal of Psychology, 213*(4), 205-218. doi:10.1026/0044-3409.213.4.205

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464.

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*(2), 197-216. doi:10.1037/0022-3514.85.2.197

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97*(1), 17-41. doi:10.1037/a0015575
10.1037/a0015575.supp (Supplemental)

Grenier, S., Barrette, A.-M., & Ladouceur, R. (2005). Intolerance of Uncertainty and Intolerance of Ambiguity: Similarities and differences. *Personality and Individual Differences, 39*(3), 593-600. doi:http://dx.doi.org/10.1016/j.paid.2005.02.014

Griffin, Z. M. (1999). Frequency of meaning use for ambiguous and unambiguous words. *Behavior Research Methods, Instruments, & Computers, 31*(3), 520-530. doi:10.3758/BF03200731

Gullvåg, I., & Næss, A. (1996). Vagueness and Ambiguity. In M. Dascal, D. Gerhardus, K. Lorenz, & G. Meggle (Eds.), *Sprachphilosophie.Ein internationales Handbuch zeitgenössischer Forschung. 2. Halbband* (pp. 1407-1417). Berlin; New York: de Gruyter.

Hahn, A., Judd, C. M., Hirsh, H. K., & Blair, I. V. (2014). Awareness of implicit attitudes. *Journal of Experimental Psychology: General, 143*(3), 1369-1392. doi:10.1037/a0035028

Halberstadt, J., & Winkielman, P. (2013). When good blends go bad: How fluency can explain when we like and dislike ambiguity. In C. Unkelbach & R. Greifender (Eds.), *The experience of thinking: How the fluency of mental processes influences cognition and behaviour.* (pp. 133-150). New York, NY: Psychology Press.

Han, H. A., Olson, M. A., & Fazio, R. H. (2006). The influence of experimentally created extrapersonal associations on the Implicit Association Test. *Journal of Experimental Social Psychology, 42*(3), 259-272. doi:https://doi.org/10.1016/j.jesp.2005.04.006

Hauk, O., & Pulvermüller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology, 115*(5), 1090-1103. doi:https://doi.org/10.1016/j.clinph.2003.12.020

Herman, J. L., Stevens, M. J., Bird, A., Mendenhall, M., & Oddou, G. (2010). The Tolerance for Ambiguity Scale: Towards a more refined measure for international management research. *International Journal of Intercultural Relations, 34*(1), 58-65. doi:http://dx.doi.org/10.1016/j.ijintrel.2009.09.004

Hermans, D., De Houwer, J., & Eelen, P. (2001). A time course analysis of the affective priming effect. *Cognition and Emotion, 15*(2), 143-165. doi:10.1080/02699930125768

Hermans, D., Houwer, J. D., & Eelen, P. (1994). The affective priming effect: Automatic activation of evaluative information in memory. *Cognition and Emotion, 8*(6), 515-533. doi:10.1080/02699939408408957

Herring, D. R., White, K. R., Jabeen, L. N., Hinojos, M., Terrazas, G., Reyes, S. M., . . . Crites, S. L., Jr. (2013). On the automatic activation of attitudes: A quarter century of evaluative priming research. *Psychological Bulletin, 139*(5), 1062-1089. doi:10.1037/a0031309

Heycke, T., Gehrmann, S., Haaf, J. M., & Stahl, C. (2018). Of two minds or one? A registered replication of Rydell et al. (2006). *Cognition and Emotion, 32*(8), 1708-1727. doi:10.1080/02699931.2018.1429389

Higgins, E. T., & Rholes, W. S. (1978). “Saying is believing”: Effects of message modification on memory and liking for the person described. *Journal of Experimental Social Psychology, 14*(4), 363-378. doi:https://doi.org/10.1016/0022-1031(78)90032-X

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology, 13*(2), 141-154. doi:10.1016/S0022-1031(77)80007-3

Hill, B. D., Foster, J. D., Elliott, E. M., Shelton, J. T., McCain, J., & Gouvier, W. D. (2013). Need for cognition is related to higher general intelligence, fluid intelligence, and crystallized intelligence, but not working memory. *Journal of Research in Personality, 47*(1), 22-25. doi:https://doi.org/10.1016/j.jrp.2012.11.001

Hinojosa, J. A., Carretié, L., Méndez-Bértolo, C., Míguez, A., & Pozo, M. A. (2009). Arousal contributions to affective priming: Electrophysiological correlates. *Emotion, 9*(2), 164-171. doi:10.1037/a0014680

Hock, M., & Krohne, H. W. (2004). Coping With Threat and Memory for Ambiguous Information: Testing the Repressive Discontinuity Hypothesis. *Emotion, 4*(1), 65-86. doi:10.1037/1528-3542.4.1.65

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A Meta-Analysis on the Correlation Between the Implicit Association Test and Explicit Self-Report Measures. *Personality and Social Psychology Bulletin, 31*(10), 1369-1385. doi:10.1177/0146167205275613

Holley-Wilcox, P., & Blank, M. A. (1980). Evidence for Multiple Access in the Processing of Isolated Words. *Journal of Experimental Psychology: Human Perception and Performance, 6*(1), 75-84. doi:10.1037/h0078139

Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in Social Categorization: The role of prejudice and facial affect in race categorization. *Psychological Science, 15*(5), 342-345. doi:10.1111/j.0956-7976.2004.00680.x

Hughes, S., Barnes-Holmes, D., Van Dessel, P., de Almeida, J. H., Stewart, I., & De Houwer, J. (2018). On the symbolic generalization of likes and dislikes. *Journal of Experimental Social Psychology, 79*, 365-377. doi:https://doi.org/10.1016/j.jesp.2018.09.002

Hummel, J. E. (2010). Symbolic Versus Associative Learning. *Cognitive Science, 34*(6), 958-965. doi:https://doi.org/10.1111/j.1551-6709.2010.01096.x

Hussey, I., & De Houwer, J. (2018). Implicit association test as an analogical learning task. *Experimental Psychology, 65*(5), 272-285. doi:10.1027/1618-3169/a000416

Hutchings, P. B., & Haddock, G. (2008). Look black in anger: The role of implicit prejudice in the categorization and perceived emotional intensity of racially ambiguous faces. *Journal of Experimental Social Psychology, 44*(5), 1418-1420. doi:10.1016/j.jesp.2008.05.002

Hütter, M., & Rothermund, K. (2020). Automatic processes in evaluative learning. *Cognition and Emotion, 34*(1), 1-20. doi:10.1080/02699931.2019.1709315

Ihmels, M., Freytag, P., Fiedler, K., & Alexopoulos, T. (2016). Relational integrativity of prime-target pairs moderates congruity effects in evaluative priming. *Memory & Cognition, 44*(4), 565-579. doi:10.3758/s13421-015-0581-8

Ishida, T. (2019). The Effects of Meaning Dominance in the Time-Course of Activation of L2 Lexical Ambiguity Processing. *Journal of Psycholinguistic Research, 48*(6), 1269-1284. doi:10.1007/s10936-019-09657-8

Itkes, O., Kimchi, R., Haj-Ali, H., Shapiro, A., & Kron, A. (2017). Dissociating affective and semantic valence. *Journal of Experimental Psychology: General, 146*(7), 924-942. doi:10.1037/xge0000291

Jach, H. K., & Smillie, L. D. (2019). To fear or fly to the unknown: Tolerance for ambiguity and Big Five personality traits. *Journal of Research in Personality, 79*, 67-78. doi:https://doi.org/10.1016/j.jrp.2019.02.003

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika, 35*(4), 401-415. doi:10.1007/BF02291817

Kanske, P., & Kotz, S. A. (2010). Leipzig Affective Norms for German: A reliability study. *Behavior Research Methods, 42*(4), 987-991. doi:10.3758/brm.42.4.987

Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology, 81*(5), 774-788. doi:10.1037/0022-3514.81.5.774

Karpinski, A., & Steinman, R. B. (2006). The Single Category Implicit Association Test as a measure of implicit social cognition. *Journal of Personality and Social Psychology, 91*(1), 16-32. doi:10.1037/0022-3514.91.1.16

Karpinski, A., Steinman, R. B., & Hilton, J. L. (2005). Attitude Importance as a Moderator of the Relationship Between Implicit and Explicit Attitude Measures. *Personality and Social Psychology Bulletin, 31*(7), 949-962. doi:10.1177/0146167204273007

Keefe, R. (2000). *Theories of Vagueness* (1. publ ed.). Cambridge [u.a.]: Cambridge Univ. Press.

Kennedy, C. (2011). Ambiguity and Vagueness. In C. Maienborn, K. Heusinger, & P. Portner (Eds.), *Semantics.An International Handbook of Natural Language Meaning* (pp. 507-534). Berlin: De Gruyter Mouton.

Kenny, D. T., & Ginsberg, R. (1958). The specificity of intolerance of ambiguity measures. *The Journal of Abnormal and Social Psychology, 56*(3), 300-304. doi:10.1037/h0041673

Kirschner, H., Hilbert, K., Hoyer, J., Lueken, U., & Beesdo-Baum, K. (2016). Psychophsyiological reactivity during uncertainty and ambiguity processing in high and low worriers. [Psychophysiologische Reaktivität während der Verarbeitung von Ungewissheit und Mehrdeutigkeit bei Menschen die sich viel und wenig sorgen]. *Journal of Behavior Therapy and Experimental Psychiatry, 50*, 97-105.

Kischkel, K.-H. (1984). Eine Skala zur Erfassung von Ambiguitätstoleranz. [Scale for assessing ambiguity tolerance]. *Diagnostica*, *30*(2), 144-154.

Klauer, K. C. (1997). Affective Priming. *European Review of Social Psychology, 8*(1), 67-103. doi:10.1080/14792779643000083

Klauer, K. C., & Musch, J. (2002). Goal-Dependent and Goal-Independent Effects of Irrelevant Evaluations. *Personality and Social Psychology Bulletin, 28*(6), 802-814. doi:10.1177/0146167202289009

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

Klauer, K. C., Musch, J., & Eder, A. B. (2005). Priming of semantic classifications: Late and response related, or earlier and more central? *Psychonomic Bulletin & Review, 12*(5), 897-903. doi:10.3758/BF03196783

Klauer, K. C., Teige-Mocigemba, S., & Spruyt, A. (2009). Contrast effects in spontaneous evaluations: A psychophysical account. *Journal of Personality and Social Psychology, 96*(2), 265-287. doi:10.1037/a0013248

Klauer, K. C., Voss, A., Schmitz, F., & Teige-Mocigemba, S. (2007). Process components of the Implicit Association Test: A diffusion-model analysis. *Journal of Personality and Social Psychology, 93*(3), 353-368. doi:10.1037/0022-3514.93.3.353

Klein, D. E., & Murphy, G. L. (2001). The Representation of Polysemous Words. *Journal of Memory and Language, 45*(2), 259-282. doi:http://dx.doi.org/10.1006/jmla.2001.2779

Klein, D. E., & Murphy, G. L. (2002). Paper has been my ruin: conceptual relations of polysemous senses. *Journal of Memory and Language, 47*(4), 548-570. doi:http://dx.doi.org/10.1016/S0749-596X(02)00020-7

Klepousniotou, E. (2002). The Processing of Lexical Ambiguity: Homonymy and Polysemy in the Mental Lexicon. *Brain and Language, 81*(1–3), 205-223. doi:http://dx.doi.org/10.1006/brln.2001.2518

Klepousniotou, E., & Baum, S. R. (2007). Disambiguating the ambiguity advantage effect in word recognition: An advantage for polysemous but not homonymous words. *Journal of Neurolinguistics, 20*(1), 1-24. doi:http://dx.doi.org/10.1016/j.jneuroling.2006.02.001

Klepousniotou, E., Pike, G. B., Steinhauer, K., & Gracco, V. (2012). Not all ambiguous words are created equal: An EEG investigation of homonymy and polysemy. *Brain and Language, 123*(1), 11-21. doi:10.1016/j.bandl.2012.06.007

Klepousniotou, E., Titone, D., & Romero, C. (2008). Making sense of word senses: The comprehension of polysemy depends on sense overlap. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(6), 1534-1543. doi:10.1037/a0013012

Klinger, M. R., Burton, P. C., & Pitts, G. S. (2000). Mechanisms of unconscious priming: I Response competition, not spreading activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(2), 441-455. doi:10.1037/0278-7393.26.2.441

Koch, S., Holland, R. W., & van Knippenberg, A. (2009). Lateralisation of diffuse positive and negative affect: Ascribing valence to ambiguous stimuli. *Cognition and Emotion, 23*(3), 587-598. doi:10.1080/02699930802028456

Kornblum, S., & Lee, J.-W. (1995). Stimulus-response compatibility with relevant and irrelevant stimulus dimensions that do and do not overlap with the response. *Journal of Experimental Psychology: Human Perception and Performance, 21*(4), 855-875. doi:10.1037/0096-1523.21.4.855

Krohne, H. W., Egloff, B., Kohlmann, C.-W., & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der "Positive and Negative Affect Schedule" (PANAS). *Diagnostica, 42*, 139-156.

Lahl, O., Göritz, A. S., Pietrowsky, R., & Rosenberg, J. (2009). Using the World-Wide Web to obtain large-scale word norms: 190,212 ratings on a set of 2,654 German nouns. *Behavior Research Methods, 41*(1), 13-19. doi:10.3758/brm.41.1.13

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

Lang, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist, 50*(5), 372-385. doi:10.1037/0003-066X.50.5.372

Lauriola, M. m. l. u. i., Foschi, R., Mosca, O., & Weller, J. (2016). Attitude Toward Ambiguity. *Assessment, 23*(3), 353-373. doi:10.1177/1073191115577188

Lawson, C., MacLeod, C., & Hammond, G. (2002). Interpretation revealed in the blink of an eye: Depressive bias in the resolution of ambiguity. *Journal of Abnormal Psychology, 111*(2), 321-328. doi:10.1037/0021-843X.111.2.321

Leichsenring, F., Steuernagel, E., Steuernagel, J., & Meyer, H. A. (2007). Tolerance of Ambiguity: Text Analytic vs Self-Report Measures in Two Nonclinical Groups. *Perceptual and Motor Skills, 104*(3), 855-869. doi:10.2466/pms.104.3.855-869

Livingston, R. W., & Brewer, M. B. (2002). What are we really priming? Cue-based versus category-based processing of facial stimuli. *Journal of Personality and Social Psychology, 82*(1), 5-18. doi:10.1037/0022-3514.82.1.5

Löbner, S. (2015). *Semantik* (2nd revised, updated and expanded edition ed.). Berlin/Boston: de Gruyter.

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods, 49*(4), 1494-1502. doi:10.3758/s13428-016-0809-y

Mac Donald, A. P. (1970). Revised Scale for Ambiguity Tolerance: Reliability and Validity. *Psychological Reports, 26*(3), 791-798. doi:10.2466/pr0.1970.26.3.791

MacLeod, C., & Cohen, I. L. (1993). Anxiety and the interpretation of ambiguity: A text comprehension study. *Journal of Abnormal Psychology, 102*(2), 238-247. doi:10.1037/0021-843X.102.2.238

Maio, G. R., Haddock, G., & Verplanken, B. (2018). *The psychology of attitudes and attitude change*: Sage Publications Limited.

Marsh, K. L., Johnson, B. T., & Scott-Sheldon, L. A. J. (2001). Heart Versus Reason in Condom Use: Implicit Versus Explicit Attitudinal Predictors of Sexual Behavior. *Experimental Psychology, 48*(2), 161-175. doi:10.1026//0949-3946.48.2.161

Martin, J. G., & Westie, F. R. (1959). The Tolerant Personality. *American Sociological Review, 24*(4), 521-528. doi:10.2307/2089539

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 3-23. doi:10.1037/0278-7393.21.1.3

McLain, D. L. (1993). The Mstat-I: A New Measure of an Individual'S Tolerance for Ambiguity. *Educational and Psychological Measurement, 53*(1), 183-189.

McLain, D. L. (2009). Evidence of the Properties of an Ambiguity Tolerance Measure: The Multiple Stimulus Types Ambiguity Tolerance Scale–II (MSTAT–II). *Psychological Reports, 105*(3), 975-988. doi:10.2466/pr0.105.3.975-988

Mehrabian, A. (1970). A semantic space for nonverbal behavior. *Journal of Consulting and Clinical Psychology, 35*(2), 248-257. doi:10.1037/h0030083

Meissner, F., & Rothermund, K. (2013). Estimating the contributions of associations and recoding in the Implicit Association Test: The ReAL model for the IAT. *Journal of Personality and Social Psychology, 104*(1), 45-69. doi:10.1037/a0030734

10.1037/a0030734.supp (Supplemental)

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*(2), 227-234. doi:10.1037/h0031564

Mierke, J., & Klauer, K. C. (2003). Method-Specific Variance in the Implicit Association Test. *Journal of Personality and Social Psychology, 85*(6), 1180-1192. doi:10.1037/0022-3514.85.6.1180

Millis, M. L., & Button, S. B. (1989). The effect of polysemy on lexical decision time: Now you see it, now you don't. *Memory & Cognition, 17*(2), 141-147. doi:10.3758/BF03197064

Mitchell, J. P., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology: General, 132*(3), 455-469. doi:10.1037/0096-3445.132.3.455

Mitchell, T., Bull, R., & Cleland, A. A. (2012). Implicit response-irrelevant number information triggers the SNARC effect: Evidence using a neural overlap paradigm. *Quarterly Journal of Experimental Psychology, 65*(10), 1945-1961. doi:10.1080/17470218.2012.673631

Moors, A. (2016). Automaticity: Componential, Causal, and Mechanistic Explanations. *Annual Review of Psychology, 67*(1), 263-287. doi:10.1146/annurev-psych-122414-033550

Moors, A., & De Houwer, J. (2006). Automaticity: A Theoretical and Conceptual Analysis. *Psychological Bulletin, 132*(2), 297-326. doi:10.1037/0033-2909.132.2.297

Moors, A., Spruyt, A., & De Houwer, J. (2010). In search of a measure that qualifies as implicit: Recommendations based on a decompositional view of automaticity. In *Handbook of implicit social cognition: Measurement, theory, and applications.* (pp. 19-37). New York, NY, US: The Guilford Press.

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology, 64*(5), 723-739. doi:10.1037/0022-3514.64.5.723

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106*(3), 226-254. doi:10.1037/0096-3445.106.3.226

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

Nievas, F., & Justicia, F. (2004). A cross-sectional study about meaning access processes for homographs. *Cognitive Development, 19*(1), 95-109. doi:https://doi.org/10.1016/j.cogdev.2003.11.002

Nievas, F., Justicia, F., Cañas, J. J., & Bajo, M. T. (2005). Lexical Processing of Ambiguous Words: Dominance or Associative Strength? *The Spanish Journal of Psychology, 8*(2), 157-179. doi:10.1017/S1138741600005059

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231-259. doi:10.1037/0033-295X.84.3.231

Nohlen, H. U., van Harreveld, F., Rotteveel, M., Barends, A. J., & Larsen, J. T. (2016). Affective responses to ambivalence are context-dependent: A facial EMG study on the role of inconsistency and evaluative context in shaping affective responses to ambivalence. *Journal of Experimental Social Psychology, 65*, 42-51. doi:https://doi.org/10.1016/j.jesp.2016.02.001

Norton, R. W. (1975). Measurement of Ambiguity Tolerance. *Journal of Personality Assessment, 39*(6), 607.

Nosek, B. A. (2005). Moderators of the Relationship Between Implicit and Explicit Evaluation. *Journal of Experimental Psychology: General, 134*(4), 565-584. doi:10.1037/0096-3445.134.4.565

Nosek, B. A., & Banaji, M. R. (2001). The Go/No-Go Association Task. *Social Cognition, 19*(6), 625-666. doi:10.1521/soco.19.6.625.20886

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and Using the Implicit Association Test: II. Method Variables and Construct Validity. *Personality and Social Psychology Bulletin, 31*(2), 166-180. doi:10.1177/0146167204271418

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at Age 7: A Methodological and Conceptual Review. In J. A. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes.* (pp. 265-292). New York, NY: Psychology Press.

Nosek, B. A., & Hansen, J. J. (2008). The associations in our heads belong to us: Searching for attitudes and knowledge in implicit evaluation. *Cognition and Emotion, 22*(4), 553-594. doi:10.1080/02699930701438186

Nosek, B. A., & Smyth, F. L. (2007). A multitrait-multimethod validation of the Implicit Association Test: Implicit and explicit attitudes are related but distinct constructs. *Experimental Psychology, 54*(1), 14-29. doi:10.1027/1618-3169.54.1.14

O'Connor, P. (1952). Ethnocentrism, 'intolerance of ambiguity,' and abstract reasoning ability. *The Journal of Abnormal and Social Psychology, 47*(2, Suppl), 526-530. doi:10.1037/h0056142

Olson, M. A., & Fazio, R. H. (2003). Relations between implicit measures of prejudice: What are we measuring? *Psychological Science, 14*(6), 636-639. doi:10.1046/j.0956-7976.2003.psci_1477.x

Olson, M. A., & Fazio, R. H. (2004). Reducing the Influence of Extrapersonal Associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality and Social Psychology, 86*(5), 653-667. doi:10.1037/0022-3514.86.5.653

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition, 9*(3), 225-236. doi:10.3758/BF03196957

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Oxford, England: Univer. Illinois Press.

Ostendorf, F., & Angleitner, A. (2004). *Neo-PI-R: Neo-Persönlichkeitsinventar nach Costa und McCrae. Revidierte Fassung*. Göttingen: Hogrefe.

Owen, H. E., Halberstadt, J., Carr, E. W., & Winkielman, P. (2016). Johnny Depp, Reconsidered: How Category-Relative Processing Fluency Determines the Appeal of Gender Ambiguity. *PLoS ONE, 11*(2), e0146328. doi:10.1371/journal.pone.0146328

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89*(3), 277-293. doi:10.1037/0022-3514.89.3.277

Perugini, M., Richetin, J., & Zogmaister, C. (2010). Prediction of behavior. In *Handbook of implicit social cognition: Measurement, theory, and applications* (Vol. 10, pp. 255-278).

Pexman, P. M., Hino, Y., & Lupker, S. J. (2004). Semantic Ambiguity and the Process of Generating Meaning From Print. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(6), 1252-1270. doi:10.1037/0278-7393.30.6.1252

Pierucci, S., Echterhoff, G., Marchal, C., & Klein, O. (2014). Creating shared reality about ambiguous sexual harassment: The role of stimulus ambiguity in audience-tuning effects on memory. *Journal of Applied Research in Memory and Cognition, 3*(4), 300-306. doi:10.1016/j.jarmac.2014.07.007

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology, 17*(3), 715-734. doi:10.1017/S0954579405050340

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology, 61*(3), 380-391. doi:10.1037/0022-3514.61.3.380

R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rammstedt, B., & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K). *Diagnostica, 51*(4), 195-206. doi:10.1026/0012-1924.51.4.195

Raphael, D., Moss, S., & Cross, W. (1978). Budner's Intolerance of Ambiguity: A Note concerning Intelligence. *Psychological Reports, 43*(2), 624-626. doi:10.2466/pr0.1978.43.2.624

Rastle, K. (2007). Visual word recognition. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 71 -87): Oxford University Press.

Rawlings, D., Barrantes i Vidal, N., & Furnham, A. (2000). Personality and aesthetic preference in Spain and England: two studies relating sensation seeking and openness to experience to liking for paintings and music. *European Journal of Personality, 14*(6), 553-576.

Rayner, K., Pacht, J. M., & Duffy, S. A. (1994). Effects of Prior Encounter and Global Discourse Bias on the Processing of Lexically Ambiguous Words: Evidence From Eye Fixations. *Journal of Memory and Language, 33*(4), 527-544. doi:http://dx.doi.org/10.1006/jmla.1994.1025

Reed, S. K. (2016). A Taxonomic Analysis of Abstraction. *Perspectives on Psychological Science, 11*(6), 817-837. doi:10.1177/1745691616646304

Reis, J. (1996). *Inventar zur Messung der Ambiguitätstoleranz (IMA)*. Heidelberg: Roland Asanger.

Rettie, H., & Daniels, J. (2020). Coping and tolerance of uncertainty: Predictors and mediators of mental health during the COVID-19 pandemic. *American Psychologist*. doi:10.1037/amp0000710

10.1037/amp0000710.supp (Supplemental)

Rimmon, S. (1977). *The Concept of Ambiguity*. Chicago: University of Chicago Press.

Rosenberg, M. J., Hovland, C. I., McGuire, W. J., Abelson, R. P., & Brehm, J. W. (1960). *Attitude organization and change: An analysis of consistency among attitude components (Yales studies in attitude and communication), Vol III*. Oxford: Yale Univer. Press.

Rothermund, K., Teige-Mocigemba, S., Gast, A., & Wentura, D. (2009). Minimizing the influence of recoding in the Implicit Association Test: The Recoding-Free Implicit Association Test (IAT-RF). *The Quarterly Journal of Experimental Psychology, 62*(1), 84-98. doi:10.1080/17470210701822975

Rothermund, K., & Wentura, D. (2004). Underlying Processes in the Implicit Association Test: Dissociating Salience From Associations. *Journal of Experimental Psychology: General, 133*(2), 139-165. doi:10.1037/0096-3445.133.2.139

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 9*(5), 487-494. doi:http://dx.doi.org/10.1016/S0022-5371(70)80091-3

Rudman, L. A., & Kilianski, S. E. (2000). Implicit and Explicit Attitudes Toward Female Authority. *Personality and Social Psychology Bulletin, 26*(11), 1315-1328. doi:10.1177/0146167200263001

Rudman, L. A., & Lee, M. R. (2002). Implicit and explicit consequences of exposure to violent and misogynous rap music. *Group Processes & Intergroup Relations, 5*(2), 133-150. doi:10.1177/1368430202005002541

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178. doi:10.1037/h0077714

Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition and Emotion, 23*(7), 1259-1283. doi:10.1080/02699930902809375

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology, 91*(6), 995-1008. doi:10.1037/0022-3514.91.6.995

Rydell, S. T. (1966). Tolerance of Ambiguity and Semantic Differential Ratings. *Psychological Reports, 19*(3_suppl), 1303-1312. doi:10.2466/pr0.1966.19.3f.1303

Rydell, S. T., & Rosen, E. (1966). Measurement and Some Correlates of Need-Cognition. *Psychological Reports, 19*(1), 139-165. doi:10.2466/pr0.1966.19.1.139

Schmidtke, D. S., Schröder, T., Jacobs, A. M., & Conrad, M. (2014). ANGST: Affective norms for German sentiment terms, derived from the affective norms for English words. *Behavior Research Methods, 46*(4), 1108-1118. doi:10.3758/s13428-013-0426-y

Schmukle, S. C., & Egloff, B. (2005). A Latent State-Trait Analysis of Implicit and Explicit Personality Measures. *European Journal of Psychological Assessment, 21*(2), 100-107. doi:10.1027/1015-5759.21.2.100

Schvaneveldt, R. W., Meyer, D. E., & Becker, C. A. (1976). Lexical ambiguity, semantic context, and visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 2*(2), 243-256. doi:10.1037/0096-1523.2.2.243

Schwarz, N. (2007). Attitude Construction: Evaluation in Context. *Social Cognition, 25*(5), 638-656. doi:10.1521/soco.2007.25.5.638

Scott, W. A. (1966). Brief report: Measures of cognitive structure. *Multivariate Behavioral Research, 1*(3), 391-395. doi:10.1207/s15327906mbr0103_9

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14*(4), 489-537. doi:10.1016/0010-0285(82)90017-2

Sherman, J. W. (2008). Controlled influences on implicit measures: Confronting the myth of process-purity and taming the cognitive monster. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures.* (pp. 391-426). New York, NY: Psychology Press.

Sherman, J. W., Gawronski, B., Gonsalkorale, K., Hugenberg, K., Allen, T. J., & Groom, C. J. (2008). The self-regulation of automatic associations and behavioral impulses. *Psychological Review, 115*(2), 314-335. doi:10.1037/0033-295X.115.2.314

Sherman, J. W., Krieglmeyer, R., & Calanchini, J. (2014). Process models require process measures. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind.* (pp. 121-138). New York, NY: The Guilford Press.

Sherman, S. J., Rose, J. S., Koch, K., Presson, C. C., & Chassin, L. (2003). Implicit and Explicit Attitudes Toward Cigarette Smoking: the Effects of Context and Motivation. *Journal of Social and Clinical Psychology, 22*(1), 13-39. doi:10.1521/jscp.22.1.13.22766

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428. doi:10.1037/0033-2909.86.2.420

Simmons, J. P., & Prentice, D. A. (2006). Pay attention! Attention to the primes increases attitude assessment accuracy. *Journal of Experimental Social Psychology, 42*(6), 784-791. doi:https://doi.org/10.1016/j.jesp.2005.12.001

Simpson, G. B. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning & Verbal Behavior, 20*(1), 120-136. doi:10.1016/S0022-5371(81)90356-X

Simpson, G. B. (1984). Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin, 96*(2), 316-340. doi:10.1037/0033-2909.96.2.316

Simpson, G. B., & Burgess, C. (1985). Activation and selection processes in the recognition of ambiguous words. *Journal of Experimental Psychology: Human Perception and Performance, 11*(1), 28-39. doi:10.1037/0096-1523.11.1.28

Simpson, G. B., & Foster, M. R. (1986). Lexical ambiguity and children's word recognition. *Developmental Psychology, 22*(2), 147-154. doi:10.1037/0012-1649.22.2.147

Simpson, G. B., & Krueger, M. A. (1991). Selective access of homograph meanings in sentence context. *Journal of Memory and Language, 30*(6), 627-643. doi:https://doi.org/10.1016/0749-596X(91)90029-J

Song, X., Chen, J., & Proctor, R. W. (2017). Role of hand dominance in mapping preferences for emotional-valence words to keypress responses. *Acta Psychologica, 180*, 33-39. doi:https://doi.org/10.1016/j.actpsy.2017.08.007

Spencer-Rodgers, J., Srivastava, S., Boucher, H. C., English, T., Paletz, S. B., & Peng, K. (2015). *The dialectical self scale. Unpublished manuscript*. California Polytechnic State University, San Luis Obispo.

Spruyt, A., De Houwer, J., Everaert, T., & Hermans, D. (2012). Unconscious semantic activation depends on feature-specific attention allocation. *Cognition, 122*(1), 91-95. doi:https://doi.org/10.1016/j.cognition.2011.08.017

Spruyt, A., De Houwer, J., Hermans, D., & Eelen, P. (2007). Affective priming of nonaffective semantic categorization responses. *Experimental Psychology, 54*(1), 44-53. doi:10.1027/1618-3169.54.1.44

Spruyt, A., Houwer, J. D., & Hermans, D. (2009). Modulation of automatic semantic priming by feature-specific attention allocation. *Journal of Memory and Language, 61*(1), 37-54. doi:https://doi.org/10.1016/j.jml.2009.03.004

Stapel, D. A., & Koomen, W. (2000). How far do we go beyond the information given? The impact of knowledge activation on interpretation and inference. *Journal of Personality and Social Psychology, 78*(1), 19-37. doi:10.1037/0022-3514.78.1.19

Stapel, D. A., & Koomen, W. (2005). When Less Is More: The Consequences of Affective Primacy for Subliminal Priming Effects. *Personality and Social Psychology Bulletin, 31*(9), 1286-1295. doi:10.1177/0146167205275616

Storbeck, J., & Clore, G. L. (2008a). Affective Arousal as Information: How Affective Arousal Influences Judgments, Learning, and Memory. *Social and Personality Psychology Compass, 2*(5), 1824-1843. doi:10.1111/j.1751-9004.2008.00138.x

Storbeck, J., & Clore, G. L. (2008b). The affective regulation of cognitive priming. *Emotion, 8*(2), 208-215. doi:10.1037/1528-3542.8.2.208

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(5), 1140-1154. doi:10.1037/0278-7393.21.5.1140

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643-662. doi:10.1037/h0054651

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*(6), 645-659. doi:https://doi.org/10.1016/S0022-5371(79)90355-4

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin, 110*(1), 67-85. doi:10.1037/0033-2909.110.1.67

Teige-Mocigemba, S., Klauer, K. C., & Sherman, J. W. (2010). A practical guide to implicit association tests and related tasks. In B. Gawronski, B. K. Payne, B. Gawronski, & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications.* (pp. 117-139). New York, NY, US: Guilford Press.

Unkelbach, C., Fiedler, K., Bayer, M., Stegmüller, M., & Danner, D. (2008). Why positive information is processed faster: The density hypothesis. *Journal of Personality and Social Psychology, 95*(1), 36-49. doi:10.1037/0022-3514.95.1.36

van Casteren, M., & Davis, M. H. (2007). Match: A program to assist in matching the conditions of factorial experiments. *Behavior Research Methods, 39*(4), 973-978. doi:10.3758/bf03192992

van Harreveld, F., Rutjens, B. T., Rotteveel, M., Nordgren, L. F., & van der Pligt, J. (2009). Ambivalence and decisional conflict as a cause of psychological discomfort: Feeling tense before jumping off the fence. *Journal of Experimental Social Psychology, 45*(1), 167-173. doi:https://doi.org/10.1016/j.jesp.2008.08.015

Van Hiel, A., Onraet, E., & De Pauw, S. (2010). The Relationship Between Social-Cultural Attitudes and Behavioral Measures of Cognitive Style: A Meta-Analytic Integration of Studies. *Journal of Personality, 78*(6), 1765-1800. doi:10.1111/j.1467-6494.2010.00669.x

Võ, M. L. H., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M. J., & Jacobs, A. M. (2009). The Berlin Affective Word List Reloaded (BAWL-R). *Behavior Research Methods, 41*(2), 534-538. doi:10.3758/brm.41.2.534

Voss, A., Rothermund, K., Gast, A., & Wentura, D. (2013). Cognitive processes in associative and categorical priming: A diffusion model analysis. *Journal of Experimental Psychology: General, 142*(2), 536-559. doi:10.1037/a0029459

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*(6), 1063-1070. doi:10.1037/0022-3514.54.6.1063

Webster, D. M., & Kruglanski, A. W. (1994a). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology, 67*(6), 1049-1062. doi:10.1037/0022-3514.67.6.1049

Webster, D. M., & Kruglanski, A. W. (1994b). Need for Closure Scale [Database record]. In. Retrieved from PsycTESTS.

Wentura, D., & Rothermund, K. (2014). Priming is not Priming is not Priming. *Social Cognition, 32*(Supplement), 47-67. doi:10.1521/soco.2014.32.supp.47

Williams, J. N. (1992). Processing polysemous words in context: Evidence for interrelated meanings. *Journal of Psycholinguistic Research, 21*(3), 193-218. doi:10.1007/BF01068072

Williams, J. N., & Colombo, L. (1995). Constraints on the range of context-independent priming from ambiguous words. *Psychological Research, 58*(1), 38-50.

Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology, 81*(5), 815-827. doi:10.1037/0022-3514.81.5.815

Wolfradt, U., Oubaid, V., Straube, E. R., Bischoff, N., & Mischo, J. (1999). Thinking styles, schizotypal traits and anomalous experiences. *Personality and Individual Differences, 27*(5), 821-830. doi:https://doi.org/10.1016/S0191-8869(99)00031-8

Wood, G., Willmes, K., Nuerk, H.-C., & Fischer, M. H. (2008). On the cognitive link between space and number: A meta-analysis of the SNARC effect. *Psychology science quarterly, 50*(4), 489.

Worthington, R. L., & Whittaker, T. A. (2006). Scale Development Research: A Content Analysis and Recommendations for Best Practices. *The Counseling Psychologist, 34*(6), 806-838. doi:10.1177/0011000006288127

Wright, N. A., & Meade, A. W. (2012). An exploration of cognitive ability contamination in the Implicit Association Test methodology. *Computers in Human Behavior, 28*(2), 393-399. doi:https://doi.org/10.1016/j.chb.2011.10.009

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*(2), 151-175. doi:10.1037/0003-066X.35.2.151

Ziegler, R. (2010). Ambiguität und Ambivalenz in der Psychologie: Begriffsverständnis und Begriffsverwendung. *Zeitschrift für Literaturwissenschaft und Linguistik, 158*, 125-171.

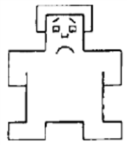
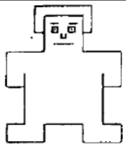
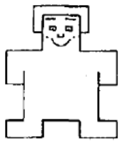
Ziegler, R., & Diehl, M. (2003). Is politician A or politician B more persuasive? recipients' source preference and the direction of biased message processing. *European Journal of Social Psychology, 33*(5), 623-637. doi:10.1002/ejsp.174

Appendix

Appendix A1

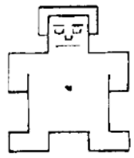
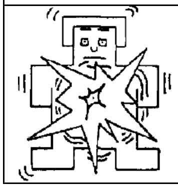
Instructions for the word rating task (valence) used in the database study

Nun werden Dir 100 verschiedene Wörter präsentiert. Du sollst jedes Wort dahingehend einschätzen, ob es für Dich eher angenehm, unangenehm oder neutral ist. Dabei geht es um Deinen ersten Eindruck. Es ist also wichtig, dass Du schnell eine Taste drückst. Mit den Zahlen 1 bis 9 kannst Du angeben, wie Du Dich beim Lesen des Wortes fühlst: von 1 (maximal unglücklich, ärgerlich, unbefriedigt, melancholisch, verzweifelt, gelangweilt) über 5 (neutrale Gefühle) bis 9 (maximal glücklich, froh, zufrieden, befriedigt, hoffnungsvoll) kannst Du alle Abstufungen vornehmen.

	Das traurige Männchen steht für negative unangenehme Gefühle bei einem Wort, z.B. Idiot.
	Das Männchen mit dem geraden Mund steht für neutrale Wörter, z.B. Antenne.
	Das lächelnde Männchen steht für positive angenehme Gefühle bei einem Wort, z.B. Ferien.

Instructions for the word rating task (arousal) used in the database study

Nun werden Dir 100 verschiedene Wörter präsentiert. Du sollst jedes Wort dahingehend einschätzen, wie aktivierend/anregend/erregend es für Dich ist. Dabei geht es um Deinen ersten Eindruck. Es ist also wichtig, dass Du schnell eine Taste drückst. Mit den Zahlen von 1 bis 9 kannst Du angeben, wie Du Dich beim Lesen des Wortes fühlst: von 1 (maximal entspannt, ruhig, träge, öde, schläfrig, unerregt) bis 9 (maximal angeregt, aufgeregt, rasend, nervös, wach, erregt) kannst Du alle Abstufungen vornehmen.

	Dieses Männchen steht für Worte die nicht aktivierend/anregend/erregend sind, wie z.B. Gremium oder Reform.
	Dieses Männchen steht für Worte die aktivierend/anregend/erregend sind, wie z.B. Gemetzel, Erektion oder Explosion.

Instructions for the word rating task (abstractness) used in the database study

„Nun werden Dir 100 verschiedene Wörter präsentiert. Du sollst jedes Wort dahingehend einschätzen, ob es für Dich eher konkret oder abstrakt ist. Dabei geht es um Deinen ersten Eindruck. Es ist also wichtig, dass Du schnell eine Taste drückst. Mit den Zahlen von 1 bis 9 kannst Du bewerten, wie konkret oder abstrakt Du das Wort findest: von 1 (sehr konkret) bis 9 (sehr abstrakt) kannst Du alle Abstufungen vornehmen.

Konkret sind all jene Dinge, die irgendwie bildlich sind, die man zum Beispiel aufmalen oder anfassen könnte (Tisch), aber auch z.B. Luft wäre konkret, man kann sie sinnlich fühlen. Abstrakt sind all jene Dinge, die nicht bildlich sind, bei denen man keine bildliche Vorstellung hat, wie bei Konzepten oder Theorien (z.B. Demokratie, Erkenntnistheorie).“

Instructions for the word meaning rating task (valence) used in the database study

„Nun werden Dir 100 verschiedene Wörter präsentiert. Die Wörter haben mehrere Bedeutungen. Eine Bedeutung wird Dir in Klammern zu jedem Wort genannt. Du sollst diese Dir genannte Bedeutung dahingehend einschätzen, ob sie für Dich eher angenehm, unangenehm oder neutral ist. Dabei geht es um Deinen ersten Eindruck. Es ist also wichtig, dass Du schnell eine Taste drückst. Mit den Zahlen 1 bis 9 kannst Du angeben, wie Du Dich beim Lesen des Wortes mit seiner Bedeutung fühlst: von 1 (maximal unglücklich, ärgerlich, unbefriedigt, melancholisch, verzweifelt, gelangweilt) über 5 (neutrale Gefühle) bis 9 (maximal glücklich, froh, zufrieden, befriedigt, hoffnungsvoll) kannst Du alle Abstufungen vornehmen.“ [added same three manikins with same explanations like in the word rating task (valence)]

Instructions for the word meaning rating task (arousal) used in the database study

„Nun werden Dir 100 verschiedene Wörter präsentiert. Die Wörter haben mehrere Bedeutungen. Eine Bedeutung wird Dir in Klammern zu jedem Wort genannt. Du sollst diese

Dir genannte Bedeutung dahingehend einschätzen, wie aktivierend/anregend/erregend sie für Dich ist. Dabei geht es um Deinen ersten Eindruck. Es ist also wichtig, dass Du schnell eine Taste drückst. Mit den Zahlen von 1 bis 9 kannst Du angeben, wie Du Dich beim Lesen des Wortes mit seiner Bedeutung fühlst: von 1 (maximal entspannt, ruhig, träge, öde, schläfrig, unerregt) bis 9 (maximal angeregt, aufgeregt, rasend, nervös, wach, erregt) kannst Du alle Abstufungen vornehmen.“ [added same two manikins with same explanations like in the word rating task (arousal)]

Instructions for the word meaning rating task (abstractness) used in the database study

„Nun werden Dir 100 verschiedene Wörter präsentiert. Die Wörter haben mehrere Bedeutungen. Eine Bedeutung wird Dir in Klammern zu jedem Wort genannt. Du sollst diese Dir genannte Bedeutung dahingehend einschätzen, ob sie für Dich eher konkret oder abstrakt ist. Mit den Zahlen von 1 bis 9 kannst Du bewerten, wie konkret oder abstrakt Du die Bedeutung des Wortes findest: von 1 (sehr konkret) bis 9 (sehr abstrakt) kannst Du alle Abstufungen vornehmen. Konkret sind all jene Dinge…“ [following: similar text like in the word rating task (abstractness)] *Instructions for the association-classification task used in the database study* „Im Folgenden bitten wir Dich, zu jeweils einem Wort Deine erste Assoziation aufzuschreiben, die Dir in den Sinn kommt. Anschließend bitten wir Dich, Deine Assoziation zu klassifizieren. Beispielsweise könnte Dir zu dem Wort Druck eine der beiden nachfolgenden Assoziationen einfallen: Zeitung oder Reifen. Diese erste Assoziation schreibst Du einfach in das Feld unter dem Wort. Anschließend präsentieren wir Dir Kategorien, wie z. B.:

[Gedrucktes Werk] [Physik: auf eine Fläche wirkende Kraft] [Andere]

„Passt Deine erste Assoziation zu einer Kategorie, dann wähle diese bitte aus, wenn nicht, wähle die Kategorie "Andere". Wir bitten Dich, diese Angaben (Assoziation und Klassifikation der Assoziation) für 100 Wörter zu machen. Dies dauert ca. 8 Minuten. Drücke auf "Weiter" um zu starten.“

Appendix A2

Summary of Variables in the NAUG Database.

	Unambiguous words			Ambiguous words			1st meaning			2nd meaning		
	M	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range
Valence	5.02	1.96	1.50 – 8.47	5.26	0.90	2.59 – 7.55	5.46	0.99	3.03 – 7.77	5.21	1.13	1.85 – 7.52
Arousal	4.28	1.38	1.95 – 7.38	3.59	0.86	2.19 – 6.00	3.37	0.84	2.15 – 6.24	3.62	0.91	2.15 – 6.33
Abstractness	3.70	1.99	1.10 – 7.26	3.13	1.51	1.45 – 7.55	2.85	1.45	1.33 – 8.05	3.63	1.57	1.43 – 6.81
Frequency	11.52	2.50	7.00 – 19.00	11.60	2.21	7.00 – 17.00						
No. of letters	5.62	1.18	3.00 – 8.00	5.58	1.60	3.00 – 13.00						

Appendix A3

Results of the association-classification task.

Word	Freq.	No. of letters	Meaning 1	Meaning 2	Freq. M1	Freq. M2	Freq. other
Akkord	15	6	Musik	Arbeitsweise	48	17	0
Anbau	11	5	Gebäude-erweiterung	Landwirtschaft	33	31	1
Angel	13	5	Fischfang	Aufhängung einer Tür	60	1	4
Annahme	11	7	Gegenstand annehmen	Hypothese	17	39	9
Aufgabe	8	7	Auftrag / Übung	Nichtfortsetzen	57	3	5
Ausbau	9	6	Entfernen von etwas	Vergrößern	1	59	4
Ball	9	4	Kugelförmiger Gegenstand	Tanzveranstaltung	49	9	6
Bank	8	4	Geldinstitut	Sitzgelegenheit	49	15	1
Blüte	13	5	Pflanze	Geld	55	8	2
Börse	10	5	Markt für Wertpapiere	Geldbörse	39	24	2
Demon-stration	10	13	Meinungs-äußerung	Verfahren um etwas darzulegen	50	10	5
Devise	12	6	Motto	Fremdwährung	34	22	9
Dichtung	14	8	Sprachliches Kunstwerk	Gegenstand zum Abdichten	28	31	4
Diele	16	5	Fußbodenbrett	Vorraum	27	30	8
Drama	11	5	Bühnenstück / Schauspiel	Trauriges, erschütterndes Geschehen	45	14	6
Eichel	14	6	Köperteil	Frucht der Eiche	16	48	1
Ente	13	4	Vogel	Falsche Pressemeldung	51	8	6
Erde	9	4	Erdboden	Planet	19	43	3
Essen	9	5	Nahrung	Stadt	58	5	2
Fahne	12	5	Flagge	Alkohol	52	11	2
Feder	12	5	Vogel	Technik	45	8	12
Flasche	11	7	Gefäß	Versager	56	7	2
Fliege	14	6	Tier	Krawattenschleife	53	7	5
Flur	12	4	Hausflur	Grundstück	61	2	2
Funken	13	6	Glühendes Teilchen	Übermittlung von Funksignalen	48	8	9
Futter	11	6	Nahrung von Tieren	Kleidung	54	5	6
Gehalt	11	6	Verdienst	Inhalt	61	4	0
Gericht	8	7	Institution	Speise	35	29	0
Geschmack	10	9	Geschmackssinn	Subjektives Urteil	49	14	1
Geschoss	14	8	Projektil / Granate	Stockwerk	30	30	4
Gipfel	10	6	Berg	Kurzfom für Gipfeltreffen	62	1	1
Grund	7	5	Boden	Ursache / Beweggrund	29	34	2

Word	Freq.	No. of letters	Meaning 1	Meaning 2	Freq. M1	Freq. M2	Freq. other
Hahn	11	4	Tier	Wasserhahn	47	17	1
Heide	12	5	Ungläubiger	Landschaft	6	52	7
Hering	13	6	Tier	Zeltbefestigung	54	9	1
Himmel	9	6	Religiöser Ort	Astronomischer Ort	12	44	9
Kamm	14	4	Haarkamm	Felskamm	54	6	5
Kanal	12	5	Wasserlauf	Rundfunk / Fernsehen	50	6	9
Kapelle	11	7	Gebäude	Musikgruppe	49	11	5
Kater	12	5	Tier	Alkohol	41	22	1
Kiefer	12	6	Köperteil	Baum	30	34	1
Kiwi	16	4	Vogel	Frucht	10	50	5
Kluft	13	5	Spalte	Kleidung	39	18	8
Knete	16	5	Knetmasse	Geld	37	25	3
Korb	12	4	Gegenstand	Ablehnende Antwort	49	9	4
Krebs	11	5	Tier	Krankheit	27	34	4
Krone	11	5	Monarchie	Zähne	38	12	15
Kunde	10	5	Käufer	Botschaft	54	5	6
Laster	13	6	Schlechte Angewohnheit	Transportwesen	25	37	3
Linie	9	5	Strich	Verkehrsstrecke	50	9	6
Lösung	8	6	Bewältigung einer Aufgabe	Chemie	50	7	8
Mandel	16	6	Samen	Lymphknoten	41	16	7
Mark	9	4	Währung	Inneres Gewebe	34	16	14
Masche	13	6	Schlinge aus Garn	Trick	43	17	5
Mast	14	4	Füttern von Nutztieren	Pfeilerähnlicher Träger	13	45	7
Maus	12	4	Tier	Computer	52	11	1
Messe	10	5	Religion	Warenschau	11	49	5
Moos	13	4	Pflanze	Geld	52	8	5
Mutter	8	6	Familie	Mechanik	57	6	2
Netz	9	4	Netzartiges Gebilde	Internet	35	24	6
Orden	12	5	Gemeinschaft	Ehrenzeichen, Abzeichen	23	41	1
Pension	12	7	Beamtenbezüge	Kleines Hotel	32	33	0
Pflaster	13	8	Straßenbelag	Heftpflaster	14	48	3
Plastik	13	7	Chemie: Material	Kunst	47	5	13
Pol	14	3	Nord- oder Südpol	Strom Aus- oder Eintrittspunkt	42	10	12
Pony	14	4	Tier	Haarschnitt	54	10	1
Preis	8	5	Geldwert	Belohnung	51	8	6
Radler	12	6	Radfahrer	Getränk	22	39	3
Reif	14	4	Ringförmiges Schmuckstück	Eiskristalle	12	30	22
Rock	10	4	Musik	Kleidung	19	45	1
Rost	13	4	Korrosion von Eisen	Gitter	56	5	3
Schale	12	6	Äußere Schicht	Gefäß	20	41	4

Word	Freq.	No. of letters	Meaning 1	Meaning 2	Freq. M1	Freq. M2	Freq. other
Schalter	12	8	Hebel / Knopf	Theke für Kundenkontakt	52	12	1
Schimmel	13	8	Pferd	Belag auf organischen Stoffen	22	42	1
Schlange	12	8	Tier	Warteschlange	52	12	1
Schloss	9	7	Vorrichtung zum Verschließen	Gebäude	16	44	5
Sender	10	6	Anlage, die Signale abstrahlt	Rundfunk-, Fernsehsender	10	50	5
Stein	10	5	Mineralische harte Masse	Kern der Steinfrucht	57	2	6
Steuer	10	6	Vorrichtung zum Steuern	Finanzabgabe	11	51	3
Stimme	9	6	Stimme des Menschen beim Sprechen / Singen	Entscheidung bei einer Wahl	59	4	2
Stock	11	5	Länglicher Gegenstand	Stockwerk	39	15	11
Strauß	12	6	Blumen	Tier	25	40	0
Strich	10	6	Linie	Prostitution	45	13	6
Strom	9	5	Fluss	Elektrizität	7	56	2
Tau	16	3	Feuchtigkeit	Starkes Seil	41	22	2
Ton	10	3	Akustik	Material für Töpferwaren	48	14	2
Tor	8	3	Durchgang	Narr	36	2	27
Veilchen	13	8	Blumen	Bluterguss am Auge	51	13	1
Verband	10	7	Wundverband	Gruppierung	49	16	0
Viertel	9	7	Wohngegend	1/4	31	30	4
Waage	12	5	Gerät zum Messen	Sternzeichen	44	17	4
Wanze	17	5	Tier	Abhörwanze	40	25	0
Wasch-lappen	16	11	Lappen zum Waschen	Feigling, Schwächling	47	17	1
Watt	13	4	Küstenstreifen	Maßeinheit	18	45	2
Weide	13	5	Pflanze	Grasbewachsenes Stück Land	13	49	3
Weizen	13	6	Getreideart	Bier	50	13	2
Werk	9	4	Produkt schöpferischer Arbeit	Industrielles Unternehmen	52	11	2
Zelle	12	5	Biologie	Gefängnis	30	30	3
Zoll	12	4	Finanzabgabe	Längeneinheit	40	16	9
Zylinder	14	8	Technik	Hut	32	28	5

Note. Freq.: frequency estimation derived from the print-based corpus of the "Wortschatz Project" from the University of Leipzig; No. Letters: Number of Letters; Freq M1(2): absolute number of selected meaning 1 (2) for this sample

Appendix A4

Mean ratings for ambiguous words and the meanings.

Word	Valence in general	Valence M1	Valence M2	Arousal in general	Arousal M1	Arousal M2	Abstractness in general	Abstractness M1	Abstractness M2
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Akkord	5.24 (2.28)	6.91 (1.54)	4.04 (2.30)	4.41 (2.42)	4.82 (2.52)	4.30 (2.57)	5.41 (3.34)	4.14 (2.33)	6.10 (2.55)
Anbau	5.76 (1.57)	5.26 (1.46)	6.11 (1.83)	2.48 (1.95)	2.29 (1.40)	3.00 (2.61)	4.14 (2.56)	2.62 (2.13)	3.52 (2.38)
Angel	5.34 (1.32)	5.14 (1.72)	4.93 (1.00)	2.30 (1.68)	2.79 (1.86)	2.24 (2.15)	1.83 (1.79)	1.52 (1.36)	2.48 (2.16)
Annahme	4.83 (1.47)	5.20 (1.45)	5.85 (1.56)	3.74 (2.41)	2.88 (1.97)	3.73 (2.30)	7.55 (2.18)	4.38 (2.62)	6.57 (2.75)
Aufgabe	4.93 (1.69)	5.06 (1.81)	3.52 (1.91)	4.67 (2.40)	3.65 (2.24)	4.27 (2.41)	6.38 (2.18)	4.43 (2.01)	6.33 (2.52)
Ausbau	5.21 (1.42)	4.69 (1.32)	5.96 (1.65)	2.44 (1.65)	3.15 (2.05)	3.09 (2.21)	5.34 (2.19)	4.86 (2.13)	4.43 (2.52)
Ball	6.21 (1.70)	6.54 (1.44)	6.07 (2.04)	3.78 (2.24)	2.76 (2.05)	4.55 (2.22)	1.55 (1.74)	1.86 (1.74)	3.57 (2.34)
Bank	4.79 (1.74)	3.86 (1.97)	6.04 (1.87)	3.93 (2.56)	3.79 (2.23)	2.45 (1.97)	2.28 (2.30)	3.71 (2.03)	1.95 (1.99)
Blüte	7.21 (1.63)	7.09 (1.56)	5.19 (2.48)	3.44 (2.55)	3.12 (2.31)	4.58 (2.31)	1.83 (2.04)	1.67 (1.46)	3.52 (2.48)
Börse	3.97 (1.52)	3.77 (1.72)	5.93 (1.47)	3.89 (2.58)	3.50 (2.54)	3.79 (2.33)	4.66 (3.21)	5.00 (2.68)	2.38 (1.88)
Demonstration	5.17 (1.97)	6.34 (1.88)	5.81 (1.96)	5.59 (2.21)	5.65 (2.12)	4.58 (2.76)	5.24 (2.52)	4.90 (2.72)	4.67 (2.76)
Devise	4.62 (1.63)	5.60 (1.50)	5.26 (1.20)	3.07 (1.94)	3.06 (2.12)	3.21 (2.52)	6.97 (2.63)	6.71 (2.37)	5.24 (2.51)
Dichtung	5.55 (1.70)	6.77 (1.66)	5.48 (1.25)	3.19 (2.30)	3.50 (2.56)	2.15 (1.44)	3.48 (2.67)	5.62 (2.73)	1.95 (1.72)
Diele	4.86 (1.22)	5.37 (1.03)	5.07 (1.38)	2.85 (2.11)	2.44 (1.83)	2.30 (1.57)	2.86 (2.12)	1.76 (1.30)	3.24 (2.23)
Drama	4.31 (1.85)	6.14 (1.75)	3.30 (2.54)	6.00 (2.27)	4.76 (2.58)	5.21 (2.58)	6.41 (2.32)	4.57 (2.46)	5.57 (2.66)
Eichel	5.59 (1.43)	5.40 (1.68)	6.00 (1.69)	2.93 (2.23)	4.50 (2.54)	3.18 (2.02)	1.97 (1.94)	2.05 (1.96)	2.10 (2.17)
Ente	6.03 (1.30)	6.17 (1.34)	3.63 (1.64)	2.63 (1.88)	2.47 (1.67)	4.45 (2.50)	1.69 (1.63)	1.33 (0.80)	5.57 (2.54)
Erde	6.41 (1.70)	6.51 (1.62)	7.19 (1.96)	3.96 (2.39)	3.15 (2.03)	4.73 (2.98)	2.17 (2.21)	1.86 (1.62)	2.29 (1.87)
Essen	7.48 (1.45)	7.31 (1.53)	5.56 (1.40)	4.52 (2.31)	4.50 (2.80)	2.48 (1.97)	1.97 (1.82)	2.33 (2.31)	3.10 (2.32)
Fahne	4.83 (1.34)	5.14 (1.40)	2.56 (1.58)	3.04 (2.33)	3.74 (2.72)	4.61 (2.60)	1.86 (2.03)	1.90 (2.02)	4.62 (2.82)
Feder	5.72 (1.51)	6.26 (1.50)	5.04 (1.32)	3.19 (2.63)	3.00 (2.10)	2.94 (2.26)	1.79 (1.80)	1.76 (1.34)	1.76 (1.34)
Flasche	5.34 (1.47)	5.69 (1.23)	3.11 (2.39)	3.22 (2.22)	2.47 (1.54)	4.64 (2.42)	1.76 (1.96)	1.48 (0.98)	5.67 (2.46)
Fliege	4.38 (1.59)	4.11 (1.86)	5.89 (1.53)	3.96 (2.61)	2.91 (2.12)	2.82 (2.48)	1.83 (2.02)	1.52 (1.25)	1.71 (1.55)
Flur	5.45 (1.35)	5.06 (1.08)	5.15 (1.26)	2.48 (1.85)	2.38 (1.56)	2.70 (1.99)	3.17 (2.17)	1.57 (0.93)	2.76 (1.73)

Word	Valence in general	Valence M1	Valence M2	Arousal in general	Arousal M1	Arousal M2	Abstractness in general	Abstractness M1	Abstractness M2
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Funken	5.97 (1.74)	6.43 (1.77)	5.41 (1.55)	5.22 (2.50)	5.18 (2.43)	3.82 (2.44)	3.28 (2.89)	1.81 (1.57)	5.24 (2.70)
Futter	5.66 (1.59)	5.77 (1.40)	5.81 (1.52)	2.67 (1.90)	2.88 (2.09)	2.52 (2.17)	2.10 (2.08)	2.19 (1.91)	2.62 (1.88)
Gehalt	7.00 (1.79)	6.09 (2.01)	5.96 (1.60)	5.33 (2.43)	4.47 (2.40)	4.21 (2.48)	5.03 (2.78)	4.00 (2.17)	6.29 (2.67)
Gericht	5.07 (1.58)	4.57 (1.90)	7.52 (1.65)	4.48 (2.31)	3.85 (2.35)	3.76 (2.48)	4.48 (2.65)	4.57 (2.46)	2.38 (1.86)
Geschmack	6.62 (1.52)	6.94 (1.63)	6.78 (1.53)	4.63 (2.50)	4.94 (2.78)	4.45 (2.44)	6.10 (2.77)	5.14 (2.56)	6.81 (2.48)
Geschoss	3.97 (1.86)	3.17 (1.77)	5.26 (1.20)	5.30 (2.77)	6.24 (2.47)	2.61 (2.09)	3.38 (2.18)	2.43 (1.99)	3.10 (2.00)
Gipfel	6.24 (1.50)	6.97 (1.56)	4.96 (1.60)	4.85 (2.67)	4.35 (2.55)	4.03 (2.56)	2.72 (2.25)	1.81 (1.17)	6.00 (2.24)
Grund	5.17 (0.97)	5.09 (1.44)	5.85 (1.68)	2.85 (2.03)	2.50 (1.58)	3.70 (2.34)	6.14 (2.91)	2.71 (2.28)	6.29 (2.59)
Hahn	4.83 (1.58)	5.74 (1.36)	5.63 (1.18)	2.74 (1.97)	2.85 (1.76)	2.70 (2.26)	1.83 (1.81)	1.43 (1.12)	2.05 (1.75)
Heide	6.00 (1.56)	4.97 (1.98)	6.59 (1.67)	2.93 (2.25)	2.97 (2.33)	3.39 (2.68)	4.10 (2.45)	5.52 (2.52)	3.62 (1.94)
Hering	4.41 (1.27)	4.94 (1.70)	5.63 (1.24)	2.67 (1.90)	2.38 (1.65)	2.79 (2.23)	1.76 (1.94)	1.48 (1.12)	1.90 (1.51)
Himmel	7.48 (1.09)	5.17 (2.63)	6.78 (2.06)	4.04 (2.86)	3.88 (2.80)	4.42 (2.97)	4.03 (3.18)	8.05 (2.25)	5.43 (2.91)
Kamm	4.90 (1.35)	5.49 (1.29)	5.22 (1.55)	2.41 (2.04)	2.44 (1.78)	3.61 (2.63)	1.72 (1.65)	1.48 (0.93)	2.81 (2.62)
Kanal	4.62 (1.50)	5.31 (1.45)	5.04 (1.40)	3.11 (2.52)	2.91 (2.12)	3.00 (2.21)	3.07 (2.51)	2.38 (1.72)	4.57 (2.60)
Kapelle	5.28 (1.73)	5.09 (1.82)	5.85 (1.32)	3.11 (2.45)	2.91 (2.11)	3.70 (2.32)	2.14 (1.83)	1.67 (0.86)	2.86 (1.62)
Kater	5.10 (2.43)	6.40 (2.03)	2.81 (1.55)	3.81 (2.83)	3.32 (2.37)	4.33 (2.48)	2.17 (2.22)	1.52 (1.36)	5.90 (2.55)
Kiefer	5.24 (1.62)	5.00 (1.43)	6.26 (1.70)	3.56 (2.64)	2.88 (1.70)	3.39 (2.21)	2.07 (2.09)	2.19 (1.63)	1.76 (1.67)
Kiwi	5.83 (1.67)	7.09 (1.31)	7.04 (1.79)	3.67 (2.66)	3.12 (2.32)	3.58 (2.44)	1.79 (1.99)	1.38 (0.80)	1.43 (1.12)
Kluft	4.14 (1.90)	4.46 (1.58)	5.59 (1.78)	4.78 (2.39)	3.97 (2.33)	3.12 (2.04)	4.59 (2.63)	2.71 (2.08)	2.71 (1.87)
Knete	5.69 (1.39)	6.09 (1.60)	6.30 (2.15)	3.41 (2.56)	3.00 (2.36)	4.42 (2.53)	2.00 (1.98)	2.00 (1.76)	3.57 (2.38)
Korb	4.69 (1.49)	5.14 (1.65)	3.00 (1.62)	2.67 (2.09)	2.29 (1.71)	5.27 (2.24)	1.59 (1.76)	1.86 (1.68)	6.24 (2.93)
Krebs	2.76 (2.13)	5.29 (1.66)	1.85 (1.77)	5.78 (2.86)	3.29 (2.22)	6.33 (2.69)	2.48 (2.49)	1.38 (0.86)	4.38 (2.73)
Krone	6.41 (1.50)	4.43 (1.72)	3.41 (2.06)	3.37 (2.34)	3.79 (2.60)	3.55 (2.60)	1.52 (1.66)	3.43 (2.79)	2.14 (1.82)
Kunde	5.10 (1.74)	4.91 (1.52)	5.70 (1.64)	3.41 (2.15)	2.88 (1.70)	3.88 (2.01)	3.28 (2.59)	2.71 (2.08)	5.52 (2.40)
Laster	4.14 (1.55)	3.31 (1.21)	4.89 (1.45)	4.41 (2.66)	4.06 (2.31)	2.76 (1.95)	3.79 (2.85)	6.10 (2.36)	2.24 (1.76)
Linie	5.24 (1.41)	5.06 (0.87)	5.63 (1.47)	2.19 (1.59)	2.15 (1.76)	2.82 (1.98)	3.28 (2.64)	2.29 (1.55)	4.24 (2.41)
Lösung	7.55 (1.18)	6.83 (1.96)	5.19 (1.52)	4.78 (2.89)	4.79 (2.88)	3.18 (1.99)	6.55 (2.13)	4.86 (2.52)	3.90 (2.55)
Mandel	5.69 (2.02)	6.00 (1.68)	4.52 (1.76)	2.63 (2.17)	2.65 (1.95)	3.21 (2.00)	1.86 (1.94)	2.33 (2.29)	3.00 (2.21)
Mark	5.21 (1.42)	5.43 (1.42)	4.74 (1.29)	3.63 (2.45)	2.76 (1.99)	3.67 (2.62)	3.55 (2.90)	3.62 (2.38)	3.43 (2.06)

Word	Valence in general	Valence M1	Valence M2	Arousal in general	Arousal M1	Arousal M2	Abstractness in general	Abstractness M1	Abstractness M2
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Masche	4.34 (1.29)	5.17 (1.29)	4.07 (1.77)	4.07 (2.45)	3.00 (1.87)	4.88 (2.30)	2.97 (2.26)	1.95 (1.66)	5.90 (2.32)
Mast	4.93 (1.25)	3.80 (2.08)	5.19 (1.24)	2.93 (2.20)	3.97 (3.04)	2.52 (1.58)	2.21 (2.29)	4.38 (2.01)	2.76 (2.55)
Maus	5.28 (1.93)	5.94 (2.06)	5.67 (1.54)	3.15 (2.23)	2.85 (1.97)	2.67 (2.12)	1.62 (1.66)	1.62 (1.63)	1.67 (1.39)
Messe	5.31 (1.61)	4.26 (1.95)	5.63 (1.78)	3.30 (2.48)	3.62 (2.51)	3.48 (2.53)	5.17 (2.48)	5.71 (2.70)	4.10 (2.41)
Moos	5.69 (1.47)	6.14 (1.33)	5.96 (2.07)	3.44 (2.64)	2.24 (1.92)	3.85 (2.56)	1.72 (1.94)	1.43 (1.12)	3.57 (2.42)
Mutter	6.97 (2.40)	7.77 (1.73)	5.26 (1.13)	4.22 (2.74)	5.00 (2.86)	3.00 (2.21)	2.21 (2.37)	2.52 (1.63)	2.05 (1.83)
Netz	5.34 (1.37)	5.31 (1.13)	6.30 (1.96)	3.37 (2.36)	3.35 (2.07)	4.15 (2.41)	2.59 (2.40)	2.19 (1.60)	6.05 (2.52)
Orden	5.86 (1.66)	4.40 (1.91)	6.00 (2.18)	3.74 (2.21)	3.50 (2.21)	3.94 (2.37)	3.17 (2.89)	5.52 (2.16)	3.33 (2.22)
Pension	5.48 (1.72)	4.83 (2.01)	6.37 (1.55)	2.81 (2.24)	3.12 (2.04)	2.97 (2.49)	4.17 (2.95)	5.00 (2.37)	2.62 (1.72)
Pflaster	4.52 (1.43)	4.69 (1.18)	5.52 (1.87)	3.74 (2.71)	2.68 (1.77)	3.12 (1.90)	1.62 (1.59)	2.33 (2.22)	1.62 (1.28)
Plastik	3.83 (1.42)	4.00 (1.81)	5.48 (2.12)	3.52 (2.38)	3.56 (2.45)	3.18 (2.10)	2.24 (2.18)	2.57 (2.04)	2.95 (2.27)
Pol	5.34 (1.37)	5.69 (1.53)	5.59 (1.37)	3.59 (2.69)	3.03 (2.41)	3.58 (2.56)	5.59 (2.78)	2.90 (1.97)	5.14 (2.76)
Pony	5.59 (1.90)	6.46 (1.79)	5.26 (1.99)	2.74 (2.18)	3.21 (2.23)	3.15 (1.95)	1.72 (1.96)	1.48 (1.44)	2.90 (2.23)
Preis	4.86 (1.77)	4.71 (1.72)	7.00 (2.09)	4.30 (2.38)	3.91 (2.14)	5.36 (2.68)	4.66 (2.83)	5.14 (2.17)	4.33 (2.76)
Radler	5.55 (1.72)	5.49 (1.54)	5.78 (2.15)	3.44 (2.22)	3.71 (2.49)	3.64 (2.36)	2.55 (2.21)	2.05 (1.88)	1.67 (1.56)
Reif	5.59 (1.30)	5.71 (1.41)	5.67 (2.18)	3.63 (2.37)	3.21 (2.40)	4.00 (2.69)	3.66 (2.66)	2.14 (2.20)	2.52 (1.91)
Rock	6.24 (1.79)	6.83 (1.84)	6.19 (1.86)	4.70 (2.91)	5.21 (2.61)	4.09 (2.73)	2.38 (2.48)	3.95 (2.04)	1.67 (1.53)
Rost	4.14 (1.25)	4.11 (1.83)	4.93 (1.38)	3.44 (2.65)	2.65 (2.01)	2.82 (2.10)	2.03 (1.61)	2.05 (1.72)	1.86 (1.71)
Schale	5.14 (1.09)	4.89 (1.13)	5.44 (1.48)	2.22 (2.03)	2.38 (1.63)	2.45 (2.20)	1.72 (1.85)	1.95 (1.20)	1.86 (1.88)
Schalter	4.93 (1.07)	5.37 (1.52)	5.33 (1.39)	3.07 (2.18)	3.00 (2.04)	3.09 (2.32)	1.86 (1.75)	1.90 (1.48)	2.57 (1.89)
Schimmel	2.59 (1.76)	5.77 (1.90)	3.22 (2.38)	3.93 (2.59)	3.71 (2.49)	4.45 (2.63)	2.24 (2.13)	1.90 (1.55)	2.33 (2.03)
Schlange	4.41 (1.99)	4.77 (2.07)	3.48 (1.72)	4.81 (2.73)	4.56 (2.81)	3.88 (2.29)	1.90 (1.74)	1.52 (1.36)	3.71 (2.61)
Schloss	6.34 (1.29)	5.03 (1.01)	6.41 (2.00)	3.70 (2.46)	3.62 (2.36)	3.91 (2.34)	1.83 (2.05)	1.67 (1.49)	1.67 (1.28)
Sender	4.90 (1.35)	4.86 (1.09)	5.11 (1.34)	3.11 (2.22)	3.38 (2.02)	3.06 (2.36)	4.45 (3.04)	4.24 (2.74)	4.19 (2.46)
Stein	5.52 (1.40)	5.71 (1.34)	5.26 (1.91)	2.74 (2.21)	2.53 (1.93)	2.85 (2.25)	1.45 (1.64)	2.00 (1.87)	2.38 (1.47)
Steuer	3.90 (2.06)	4.46 (1.82)	2.89 (1.85)	4.30 (2.92)	3.56 (2.30)	4.33 (2.52)	5.69 (3.04)	2.76 (2.34)	5.29 (2.51)
Stimme	6.24 (1.18)	6.63 (1.57)	6.44 (1.60)	5.26 (2.74)	4.62 (2.62)	5.76 (2.37)	4.14 (2.94)	3.52 (1.94)	5.48 (2.50)
Stock	4.86 (1.48)	5.26 (1.07)	5.41 (1.08)	2.63 (2.13)	2.79 (1.77)	2.48 (1.97)	1.93 (1.81)	1.38 (0.74)	2.86 (1.62)
Strauß	5.86 (1.73)	6.71 (1.64)	5.85 (1.63)	3.07 (2.30)	4.26 (2.74)	3.21 (2.22)	1.90 (1.90)	1.48 (1.03)	1.43 (0.98)

Word	Valence in general	Valence M1	Valence M2	Arousal in general	Arousal M1	Arousal M2	Abstractness in general	Abstractness M1	Abstractness M2
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Strich	4.97 (1.24)	4.74 (1.12)	3.48 (2.33)	2.89 (2.69)	2.15 (1.60)	4.97 (2.44)	3.45 (2.73)	2.67 (2.31)	5.48 (2.25)
Strom	5.55 (1.53)	6.17 (1.60)	5.44 (1.31)	4.81 (2.27)	4.15 (2.61)	4.27 (2.44)	4.55 (3.04)	2.71 (2.61)	4.43 (2.40)
Tau	5.93 (1.41)	5.74 (1.85)	5.96 (1.60)	2.63 (2.22)	2.85 (2.30)	3.36 (2.21)	2.45 (1.99)	2.05 (1.72)	1.90 (1.70)
Ton	5.76 (1.72)	6.57 (1.58)	5.78 (1.55)	4.67 (2.62)	4.29 (2.59)	3.00 (2.25)	3.72 (3.03)	3.48 (2.25)	1.95 (1.91)
Tor	5.45 (1.59)	5.46 (1.36)	4.41 (2.00)	3.44 (2.71)	3.00 (2.13)	3.70 (2.38)	2.03 (1.88)	1.86 (1.06)	4.52 (2.50)
Veilchen	5.34 (1.74)	7.09 (1.31)	3.70 (2.16)	3.33 (2.37)	2.68 (1.89)	5.42 (2.89)	2.34 (2.16)	1.62 (1.40)	2.86 (2.15)
Verband	4.45 (1.33)	4.60 (1.44)	5.33 (1.66)	3.70 (2.71)	3.94 (2.37)	3.61 (2.36)	2.55 (2.23)	2.48 (2.46)	5.57 (2.68)
Viertel	4.97 (0.82)	5.51 (1.46)	5.19 (1.44)	2.63 (1.82)	3.03 (1.80)	2.24 (1.92)	4.76 (2.86)	3.52 (2.14)	5.57 (3.28)
Waage	4.86 (1.66)	4.80 (1.51)	6.07 (2.16)	2.89 (2.15)	2.74 (1.62)	2.27 (1.64)	1.76 (1.86)	2.05 (1.83)	6.14 (3.04)
Wanze	3.45 (1.94)	3.14 (1.63)	3.11 (2.19)	3.78 (2.78)	3.79 (2.61)	5.73 (2.35)	2.45 (2.23)	1.67 (1.24)	2.14 (1.82)
Waschlappen	4.62 (1.54)	4.86 (1.22)	2.96 (1.91)	2.30 (1.79)	2.71 (2.02)	5.21 (2.81)	1.55 (1.74)	1.71 (1.49)	4.90 (2.83)
Watt	5.03 (1.12)	6.46 (1.63)	5.19 (1.52)	3.59 (2.50)	3.12 (1.87)	2.64 (1.95)	5.10 (3.00)	2.19 (1.44)	6.24 (2.96)
Weide	6.17 (1.56)	6.40 (1.59)	6.52 (2.01)	2.89 (2.31)	2.24 (1.97)	3.09 (2.39)	2.31 (1.73)	2.00 (1.48)	2.57 (2.04)
Weizen	5.86 (1.57)	6.17 (1.46)	5.44 (2.03)	2.70 (2.22)	2.32 (1.63)	3.55 (2.29)	1.79 (1.78)	1.86 (1.46)	1.67 (1.71)
Werk	5.31 (1.63)	6.74 (1.79)	4.85 (1.56)	3.59 (2.29)	3.68 (2.27)	3.00 (2.40)	4.34 (3.04)	4.95 (2.52)	3.67 (2.58)
Zelle	4.38 (2.06)	5.74 (1.20)	2.93 (2.32)	4.48 (2.68)	3.21 (2.14)	5.00 (2.30)	2.90 (2.34)	3.33 (2.54)	2.05 (1.56)
Zoll	3.72 (1.62)	3.03 (1.74)	4.93 (1.21)	3.96 (2.85)	3.29 (2.11)	2.39 (2.14)	5.17 (3.09)	5.10 (2.17)	5.14 (3.10)
Zylinder	5.52 (1.53)	4.94 (1.45)	5.56 (1.37)	3.56 (2.56)	2.94 (2.35)	2.52 (1.94)	1.97 (1.80)	1.95 (1.32)	1.81 (1.57)

Note. M1 = meaning 1, M2 = meaning 2

Appendix A5

Mean ratings for unambiguous words.

Word	Frequency	No. of letters	Valence	Arousal	Abstractness
			M (*SD*)	*M* (*SD*)	*M* (*SD*)
Akte	13	4	4.53 (1.05)	2.77 (2.26)	1.65 (1.64)
Anfahrt	13	7	4.53 (1.27)	2.44 (1.55)	5.06 (2.37)
Angriff	9	7	2.69 (1.45)	6.36 (2.21)	4.61 (2.46)
Apfel	12	5	6.66 (1.41)	2.36 (1.68)	1.13 (0.43)
Ärger	10	5	2.69 (1.15)	6.18 (2.28)	5.87 (2.36)
Asthma	13	6	2.78 (1.34)	4.64 (2.63)	3.77 (2.29)
Ausflug	11	7	7.25 (1.39)	3.82 (2.06)	4.39 (2.58)
Baby	10	4	6.47 (2.11)	4.13 (2.04)	1.26 (0.68)
Beruf	9	5	5.88 (1.56)	3.67 (2.46)	5.10 (2.40)
Bildung	9	7	7.19 (1.42)	4.49 (2.50)	6.65 (2.40)
Blume	13	5	7.16 (1.37)	3.08 (2.03)	1.10 (0.30)
Bordell	14	7	3.16 (1.80)	5.13 (2.55)	2.61 (1.94)
Brief	10	5	6.03 (1.49)	3.21 (2.08)	1.29 (0.86)
Busen	14	5	5.41 (1.78)	3.92 (2.14)	1.81 (1.58)
Butter	11	6	5.28 (1.28)	2.05 (1.19)	1.26 (0.93)
Drogen	10	6	3.38 (1.93)	5.33 (2.67)	2.94 (2.32)
Drohung	12	7	2.16 (1.19)	6.41 (2.38)	5.65 (2.46)
Effekt	11	6	5.78 (1.16)	4.05 (2.33)	6.65 (2.23)
Eifer	13	5	6.00 (1.44)	4.26 (2.16)	6.55 (2.14)
Eiter	17	5	2.03 (1.06)	4.90 (2.84)	1.87 (1.59)
Engel	11	5	7.00 (1.70)	2.72 (2.04)	4.32 (2.90)
Fackel	14	6	5.56 (1.24)	4.13 (2.18)	1.35 (0.95)
Falle	10	5	2.22 (0.94)	5.23 (2.43)	3.77 (2.36)
Flirt	14	5	6.41 (1.70)	5.23 (2.42)	5.65 (2.58)
Fokus	10	5	5.69 (1.40)	3.36 (2.19)	6.19 (2.51)
Folter	12	6	1.50 (0.84)	7.38 (1.93)	4.71 (2.60)
Freiheit	9	8	8.47 (0.76)	5.49 (2.80)	6.84 (2.25)
Freund	9	6	8.31 (0.97)	5.18 (2.68)	3.48 (2.51)
Gebiet	9	6	5.09 (0.93)	2.38 (1.91)	5.03 (2.44)
Gefahr	8	6	2.97 (1.45)	6.18 (2.28)	5.52 (2.01)
Geilheit	18	8	4.56 (2.49)	5.13 (2.63)	6.32 (2.50)
Geiz	14	4	2.91 (1.33)	4.49 (2.29)	7.10 (1.81)
Genuss	11	6	7.66 (1.33)	4.87 (2.78)	6.13 (2.23)
Gift	12	4	2.41 (1.41)	5.79 (2.35)	3.13 (2.43)
Grab	11	4	2.59 (1.43)	5.03 (2.67)	1.52 (1.39)
Hafen	10	5	6.47 (1.14)	2.59 (1.83)	1.87 (1.50)
Hälfte	8	6	5.03 (1.03)	2.36 (1.58)	4.61 (2.58)
Harem	16	5	4.16 (1.48)	4.41 (2.22)	3.90 (2.72)
Hass	11	4	1.84 (0.92)	6.82 (2.52)	6.13 (2.29)
Held	11	4	6.88 (1.77)	4.56 (2.54)	5.19 (2.64)
Henker	15	6	2.22 (1.48)	5.33 (2.78)	2.61 (2.12)
Hilfe	7	5	6.63 (1.62)	4.69 (2.13)	5.00 (2.54)
Jugend	9	6	6.16 (1.78)	4.41 (2.05)	5.90 (2.33)

Word	Frequency	No. of letters	Valence	Arousal	Abstractness
			M (*SD*)	*M* (*SD*)	*M* (*SD*)
Ketzer	16	6	3.00 (1.80)	4.38 (2.38)	5.39 (2.84)
Knast	12	5	2.25 (1.39)	4.92 (2.43)	2.45 (2.11)
Korken	15	6	5.06 (0.84)	2.41 (1.86)	1.19 (0.65)
Krise	8	5	2.75 (1.61)	5.62 (2.61)	6.16 (2.50)
Kuss	13	4	7.97 (1.15)	6.26 (2.41)	1.77 (1.52)
Lächeln	11	7	8.25 (0.92)	4.72 (2.66)	2.45 (2.06)
Leiche	11	6	2.03 (1.51)	5.87 (2.62)	1.65 (1.54)
Licht	8	5	7.31 (1.20)	3.95 (2.34)	2.42 (1.73)
Liebe	8	5	8.44 (0.88)	6.33 (2.78)	5.23 (2.97)
Lüstling	19	8	3.63 (1.72)	4.59 (2.26)	5.45 (2.23)
Masseur	16	7	6.56 (1.68)	3.51 (2.02)	1.48 (0.77)
Meer	9	4	8.19 (1.15)	3.79 (2.68)	1.48 (1.39)
Mord	10	4	1.78 (0.97)	6.67 (2.44)	3.84 (2.48)
Notiz	14	5	4.97 (0.74)	2.69 (1.64)	2.19 (1.83)
Nutzen	10	6	5.53 (1.29)	3.03 (1.77)	6.52 (2.38)
Ofen	12	4	6.25 (1.14)	2.74 (1.73)	1.29 (1.13)
Organ	13	5	5.44 (1.44)	3.33 (1.85)	2.26 (1.69)
Orgie	16	5	4.75 (2.34)	5.97 (2.72)	5.00 (2.74)
Papier	10	6	5.41 (0.95)	2.18 (1.41)	1.42 (1.23)
Rache	12	5	2.56 (1.24)	6.15 (2.07)	6.32 (2.77)
Reife	13	5	6.06 (1.19)	3.51 (2.08)	7.03 (2.32)
Ruder	12	5	5.31 (1.18)	2.18 (1.47)	1.61 (1.80)
Sau	14	3	4.22 (1.56)	3.54 (1.97)	1.68 (1.28)
Säufer	16	6	2.03 (1.12)	5.00 (2.53)	2.87 (2.36)
Schimmel	13	8	2.31 (1.53)	4.41 (2.44)	1.87 (1.50)
Schnee	9	6	6.88 (1.70)	4.46 (2.42)	1.26 (0.73)
Schnur	14	6	4.94 (0.91)	1.95 (1.47)	1.61 (1.54)
Schrei	13	6	3.25 (1.68)	6.44 (2.21)	3.81 (2.64)
Schulden	9	8	1.88 (0.79)	5.82 (2.48)	5.03 (2.43)
Seide	14	5	6.06 (1.39)	2.69 (1.89)	1.71 (1.62)
Sex	10	3	7.38 (1.70)	6.28 (2.43)	2.81 (2.12)
Slalom	12	6	5.03 (0.86)	3.13 (1.96)	3.23 (2.26)
Sommer	8	6	7.66 (1.52)	4.44 (2.65)	2.97 (2.21)
Spende	11	6	7.03 (1.43)	3.38 (1.68)	3.48 (2.03)
Spiegel	10	7	5.25 (0.92)	3.15 (2.29)	1.26 (0.68)
Spritze	14	7	3.66 (1.54)	5.03 (2.76)	1.39 (1.48)
Stahl	11	5	4.78 (1.41)	3.08 (1.98)	1.68 (1.25)
Strand	10	6	7.88 (1.21)	3.59 (2.34)	1.68 (1.38)
Straße	10	7	5.09 (0.78)	2.72 (1.96)	1.65 (1.38)
Stress	10	6	2.38 (1.72)	6.95 (2.22)	5.29 (2.57)
Sünde	12	5	3.53 (1.88)	5.23 (2.30)	7.26 (2.00)
Talent	11	6	6.88 (1.54)	4.74 (2.20)	6.29 (2.66)
Taxi	12	4	5.00 (1.14)	2.44 (1.52)	1.42 (1.39)
Technik	9	7	5.84 (1.71)	3.21 (1.98)	4.97 (2.36)
Tisch	9	5	5.28 (0.68)	2.26 (1.68)	1.29 (1.27)
Tod	8	3	2.72 (2.13)	6.31 (2.55)	4.03 (2.93)
Treppe	12	6	5.00 (0.72)	2.64 (1.69)	1.42 (1.48)
Tumor	13	5	1.50 (1.02)	6.05 (2.71)	2.52 (1.90)

Word	Frequency	No. of letters	Valence	Arousal	Abstractness
			M (*SD*)	*M* (*SD*)	*M* (*SD*)
Verkauf	9	7	5.06 (1.16)	2.67 (1.75)	4.77 (2.51)
Wachstum	9	8	5.97 (1.51)	3.79 (2.08)	5.16 (2.57)
Wahrheit	9	8	7.81 (1.15)	4.87 (2.73)	6.77 (2.92)
Wärme	11	5	7.94 (1.13)	4.59 (2.34)	3.71 (2.40)
Weste	13	5	5.09 (1.33)	2.03 (1.44)	1.74 (1.90)
Wollust	17	7	4.91 (1.63)	4.26 (2.30)	6.16 (2.13)
Wunder	9	6	7.75 (1.46)	5.13 (2.73)	6.97 (2.85)
Zelt	11	4	5.97 (1.33)	3.10 (2.01)	1.35 (1.14)
Zustand	9	7	4.84 (0.77)	2.38 (1.41)	7.00 (2.68)

Note. Frequency: frequency estimation derived from the print-based corpus of the “Wortschatz Project” from the University of Leipzig; No. Letters: Number of Letters;

Appendix B1

Instructions for description of an ambiguous situation

Unsere Forschergruppe möchte ein Inventar zu mehrdeutigen Situationen entwickeln. Unter mehrdeutigen Situationen verstehen wir Gegebenheiten, Sachverhalte oder auch Gegenstände, mit denen wir (im Alltag) konfrontiert werden und die zwei oder mehr Betrachtungen zulassen. Dafür bitten wir Dich uns in dem untenstehenden Textfeld eine oder gerne auch mehrere konkrete Situationen mitzuteilen, in denen Du mit etwas Mehrdeutigem konfrontiert warst.

Wichtig: Mit mehrdeutig meinen wir, dass es zwei oder mehrere klar voneinander unterscheidbare Interpretationen für eine Situation gibt. Eine Situation ist also dann mehrdeutig, wenn Du nicht genau weißt, wie eine Information gemeint ist, bzw. wenn die Situation auf zwei oder mehr Arten zu verstehen ist. Solche Situation können derart stattgefunden haben, dass Du nicht wusstest, was Du zu tun hattest, bzw. was die richtige Verhaltensweise war. Es kann aber auch sein, dass Du eine solche Situation lediglich wahrgenommen hast, ohne selbst involviert gewesen zu sein. Die Situationen können aus dem Arbeitsleben, dem Studium oder der Freizeit stammen und können sprachliche Äußerungen, soziale Interaktionen oder etwas Anderes sein. Es gibt keine richtigen oder falschen Antworten.

<<Place for description>>

Instructions for description of the reaction

Auch wenn die gerade geschilderte Situation eventuell schon länger her ist, bitten wir Dich in das untenstehende Textfeld einzutragen, wie Du sie wahrgenommen hast. Als Anregungen können Dir folgende Fragen dienen:
Was waren Deine Gedanken in der Situation?
Wie hast Du Dich dabei gefühlt?
Wie bist Du mit der Situation umgegangen?

<<Place for description>>

Appendix B2

Exemplary selection of responses given to the two questions from G1

Situation	Perception/Feeling of situation
manchmal weiß ich nicht ob eine ironische aussage ironisch oder ernst gemeint sein soll, dann weiß ich nicht wie ich reagieren soll also z.b. lachen oder anderweitig drauf antworten	gedanken: ich überlege mir welche bedeutung gemeint sein könnte. häufig kann man es mit vorangegangenen informationen herausfinden oder im kontext, aber manchmal eben nicht\n\ngefühl: ich fühle mich dann nicht sehr wohl, weil ich angst habe ich könnte genau die falsche interpretation erwischen und ausgelacht werden\n\numgang: ich versuche herauszufinden was gemeint ist oder antworte nicht sofort, damit ich bedenkzeit habe oder grinse nur leicht, also eine antwort auf beide fälle
1. ein kollege zwinkert dir zu und du weißt nicht, ob es eine anmache oder ein scherz ist. 2. eine aussage, bei der man nicht weiß, ob es ironisch oder ernst gemeint ist.	ich fühle mich in mehrdeutigen situationen meist zu unsicher, als dass ich sie direkt anspreche und ignoriere sie daher.
wenn jemand im z.b. bus sich neben mir sitzt und dann nach ein paar minuten aufsteht und einen anderen sitzplatz nimmt. oder wenn es mehrere freie platze neben ein paar personen wird, sitzt sich einer fast neben mir.	ich mache mir immer sorgen, dass ich unangenehm rieche (kann eig nicht sein) oder eine negative ausstrahlung habe. \nich fühle mich dabei einerseits komisch, andererseits genieße ich meine eigene ruhe und den umher freigewordenen platz. \nin solch situtation mache ich nix.
wahtsappnachricht enthält keine mimik, rhetorik, aussprache, stimmenklang und kann häufg mehrdeutig ausgelegt werden.	ich versuche die dinge bei denen ich das gefühl habe zwischen den zeilen entwas zu lesen persönlich zu klären
jemand meldet sich nicht	jemand mag mich nicht oder will keinen kontakt
mann: ich hatte bei der geburt meiner tochter den spiegel dabei...gemeint war das magazin,nicht ein handspiegel	ich habe mich gefragt,wozu man bei einer geburt einen spiegel brauchen kann und ob der angehende vater etwa alkes ganz genau sehen wollte.alle haben nach aufklärung des missverständnisses lautstark gelacht.

Appendix B3

Itempool for exploratory factor analysis used in Study 1

Item No.	Content
1	Ich finde es gut, wenn Arbeitskollegen ihre Meinung eindeutig vertreten.
2	Witze, in denen Wortspiele vorkommen, finde ich uninteressant.
3	Menschen, die widersprüchliche Aussagen machen, finde ich unangenehm.
4	Nach einem Film mit offenem Ende habe ich immer ein leichtes Gefühl der Beklemmung.
5	Wenn ich das Geschlecht einer Person nicht zuordnen kann, weil es sich z. B. um eine maskuline Frau oder einen femininen Mann handeln könnte, stört mich das.
6	Widersprüchliche Informationen stressen mich.
7	Die Äußerungen eines Vorgesetzten sollten immer klar und eindeutig sein.
8	Mehrdeutige Situationen machen mich ein wenig nervös
9	Ich beschäftige mich gerne mit eindeutigen Sachverhalten.
10	Kunst, die viele Interpretationen zulässt, meide ich.
11 (R)	Ich finde es interessant, über verschiedene potentielle Bedeutungen einer Aussage nachzudenken.
12 (R)	Ich finde es gut, dass manche Informationen mehrere Sichtweisen zulassen.
13	Wenn ich nicht weiß, was ein Bekannter mit einer bestimmten Handlung bezwecken will, vermeide ich es, dies zu analysieren.
14 (R)	Wenn in einem Roman nicht klar ist, ob das Geschilderte tatsächlich geschehen ist oder ob es nur eine Vorstellung oder ein Traum einer Romanfigur war, finde ich das besonders spannend.
15	Gegenständliche Kunst finde ich angenehmer als abstrakte Kunst.
16	Vornamen sollten eindeutig männlich oder weiblich sein.
17	Arbeitsanweisungen, die man auf zwei Arten verstehen kann, machen mich nervös.
18 (R)	Ich mag Wortspiele
19	Bei einer Diskussion sollte am Ende ein Standpunkt klar favorisiert werden.
20 (R)	Ich mag Musik, die man auf mehrere Arten interpretieren kann.
21 (R)	Ich finde ironische Aussagen bereichern den Alltag.
22	Grundfarben ziehe ich Mischfarben vor.
23	Missverständnisse sind mir peinlich.
24	Wenn eine Arbeitsanweisung nicht eindeutig formuliert ist, versuche ich diese erst mal zu ignorieren.
25	Unklare hierarchische Strukturen bei der Arbeit finde ich eher unangenehm.
26	Ich versuche soziale Situationen, in denen Missverständnisse auftreten können, zu vermeiden.
27	Ich würde am liebsten gar nicht zur Arbeit gehen, wenn ich weiß, dass es Aufgaben gibt, die nicht klar definiert sind
28	Sozialen Situationen, die Widersprüche von Mimik, Gestik, Sprache oder Handlung aufweisen, gehe ich aus dem Weg.
29	Chatnachrichten oder E-Mails, die „zwischen den Zeilen“ Informationen enthalten, können häufig zu Problemen führen.
30	Chatnachrichten oder E-Mails, die „zwischen den Zeilen“ Informationen enthalten, beunruhigen mich.
31	Man sollte Ironie und Ernst immer unterscheiden können.
32	Wenn ich nicht weiß, ob eine Aussage ironisch oder ernst gemeint ist, versuche ich dies zu klären.
33 (R)	Gute Kunst muss viel Raum für Interpretationen lassen.

Item No.	Content
34	Wenn ich eine Situation auf zwei Arten interpretieren kann, habe ich ein großes Bedürfnis Klarheit herzustellen.
35 (R)	Ich finde viele Interpretationen einer Situation oder eines Sachverhalts erweitern den Horizont.
36 (R)	Wenn sich ein/e Freund/in länger nicht meldet, denke ich mir dabei nichts Schlimmes.
37	Filme, bei denen am Ende nicht klar ist, ob es ein Happy End gibt oder nicht, schaue ich mir kein zweites Mal an.
38	Wenn ich nicht zwischen Ironie und Ernst unterscheiden kann, stresst mich das.
39	Wenn in einem literarischen Text unklar ist, ob das Geschehen, das die Figuren erleben, real oder ein Traum ist, finde ich das nicht angenehm.
40	Einen Clown, der so geschminkt ist als sei er sowohl traurig als auch fröhlich, finde ich ein wenig beängstigend.
41 (R)	Gute Literatur muss viel Raum für Interpretationen lassen.
42	Wenn eine Zeichnung keine konkreten Gegenstände darstellt, schaue ich sie mir nicht weiter an.
43	Es ärgert mich, wenn Wegweiser mehrdeutig sind.
44	Ich will Farben immer klar benennen können.
45	Es irritiert mich, wenn sich in einer Nachricht Text und Emotions widersprechen.
46	Wenn ich mir unsicher bin, ob ich jemanden duzen oder siezen sollte, vermeide ich die direkte Anrede.
47	In sozialen Netzwerken, verwende ich nur Emoticons, die eine eindeutige Emotion ausdrücken.
48 (R)	Ich lese gerne Gedichte.
49	Wenn ich bei einem Gesprächspartner, z. B. aufgrund einer Sonnenbrille, nicht genau erkennen kann, ob er mich ansieht oder nicht, finde ich das unangenehm.
50	Ein guter Job ist einer, bei dem immer klar ist, wo was zu tun ist und wie es zu tun ist (Budner, 1962)
51	Ich fühle mich ein wenig unbehaglich im Umgang mit Menschen, wenn ich das Gefühl habe, dass ich ihr Verhalten nicht verstehen kann. (Mac Donald, 1970) ursprünglich aus Rydell-Rosen 16 Item-Skala
52 (R)	Wenn ich Arzt wäre, würde ich lieber die Ungewissheiten eines Psychiaters in Kauf nehmen wollen, als die klare und definitive Arbeit von jemandem wie einem Chirurgen oder Augen- Spezialisten. (Mac Donald, 1970) ursprünglich aus Rydell-Rosen 16 Item-Skala
53	Ich mag es wirklich nicht, wenn eine Person über sich keine klaren Antworten gibt. (Norton, 1975)
54	Es würde mich stören, wenn enge Freunde widersprüchliche Meinungen über mich hätten. (Norton, 1975)
55	Ich ertrage mehrdeutige Situationen nicht gut. (McLain, 2009)
56	Ich versuche mehrdeutige Situationen zu meiden. (McLain, 2009)
57 (R)	Ich bin mehrdeutigen Situationen gegenüber tolerant. (McLain, 2009)
58	Ich mag mehrdeutige Situationen nicht. (McLain, 2009)
59 (R)	Ich bevorzuge Situationen, die ein wenig mehrdeutig sind. (McLain, 2009)
60	Ich mag es nicht, wenn die Aussage einer Person viele verschiedene Dinge bedeuten könnte. (Webster & Kruglanski, 1994)
61	Ich fühle mich unwohl, wenn die Funktion oder Absicht von jemandem unklar für mich ist. (Webster & Kruglanski, 1994)

Note. (R) = Reversed coded item, Items 50 to 61 are taken from other scales.

Appendix B4

Classification of items (Item no) into the grid 4 (domains) x 3 (classes of evaluative responses)

	Cognitive/General evaluation	Affective	Behavioral/Motivational	Items
Social related	2	3	13	21
	11	5	26	
	19	23	28	
	21	30	46	
	29	49	47	
	36	51		
	53	54		
	60	61		
Job related	1	17	24	8
	7	25	27	
	50			
	52			
Art related	18	4	10	13
	20	15	14	
	33	39	37	
	41		42	
			44	
			48	
Domain unspecific	12	6	9	19
	16	8	22	
	31	38	32	
	35	40	34	
	45	43	56	
	57	55	59	
	58			
Items	23	19	19	61

Appendix B5

Item statistics for the 61 items of the initial item pool.

Item No	*M*	*SD*	*N*
1	3.84	0.75	753
2	0.83	0.95	753
3	3.23	1.09	754
4	2.34	1.39	754
5	1.71	1.36	755
6	3.07	1.07	752
7	3.92	0.88	753
8	2.52	1.06	751
9	3.04	0.99	749
10	1.45	1.15	754
11	1.47	1.16	752
12	1.60	1.04	750
13	1.59	1.05	749
14	2.18	1.29	752
15	2.73	1.32	752
16	2.17	1.52	754
17	2.96	1.10	752
18	0.91	1.00	753
19	2.28	1.25	753
20	1.70	1.09	750
21	0.98	1.01	753
22	1.64	1.25	750
23	3.08	1.10	754
24	1.74	1.15	755
25	2.78	1.21	755
26	2.80	1.19	753
27	1.92	1.22	754
28	2.26	1.08	750
29	3.55	0.99	754
30	2.39	1.22	753
31	2.94	1.22	754
32	3.16	1.03	752
33	1.90	1.24	749
34	3.20	1.01	753
35	1.45	1.02	751
36	1.85	1.33	755
37	1.61	1.36	752
38	2.55	1.23	754
39	2.03	1.35	751
40	2.30	1.51	751
41	2.00	1.23	753

Item No	*M*	*SD*	*N*
42	1.60	1.26	750
43	3.78	1.05	755
44	1.99	1.32	753
45	3.10	1.19	751
46	3.74	1.24	755
47	2.50	1.35	748
48	2.64	1.54	753
49	3.12	1.28	753
50	2.67	1.24	754
51	3.31	1.04	753
52	3.29	1.32	753
53	3.16	1.12	748
54	2.82	1.28	753
55	2.06	1.04	750
56	2.09	1.05	749
57	1.72	0.87	751
58	2.20	1.02	748
59	2.65	0.77	750
60	2.77	1.05	751
61	3.14	0.99	744

Note. Apart from the first item all items had an observed range from 0 to 5. The first item from 1 to 5.

Appendix B6

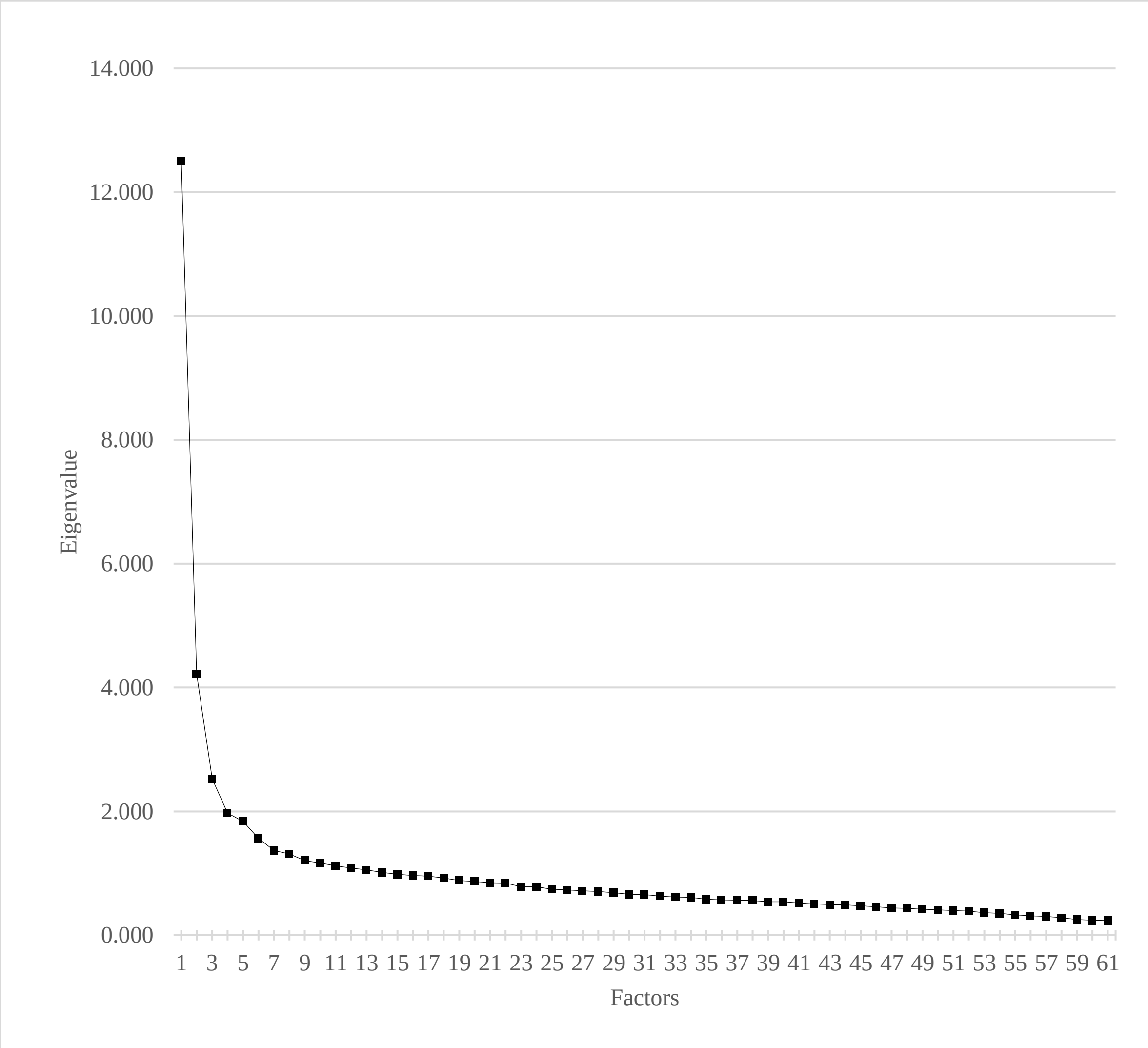


Figure 21. Scree plot of 61 factors of step 2

Appendix B7

Factor loadings for 3 factor solution after varimax rotation.

Item no	Factors		
	1	2	3
55	**.72**	.17	.16
56	**.71**	.25	.14
58	**.66**	.33	.23
8	**.64**	.09	.25
28	**.58**	.07	.16
30	**.56**	.11	.30
57	**.55**	.29	.01
38	**.53**	.05	.28
26	**.52**	.10	.21
60	**.49**	.28	.39
17	**.48**	.04	.35
51	**.47**	-.02	.34
6	**.47**	.08	**.42**
27	**.46**	.13	.14
23	**.45**	.05	.27
45	.38	.12	.33
4	.37	.22	.16
21	.35	.27	-.19
50	.34	.27	.32
40	.32	.10	.08
24	.32	.09	-.01
44	.30	.23	.15
49	.28	-.03	.14
36	.28	.09	.02
46	.22	-.03	.15
10	.13	**.65**	.10
41	-.05	**.64**	.11
33	-.05	**.62**	.14
11	.15	**.61**	-.07
35	.15	**.60**	.03
14	.01	**.57**	.19
20	.09	**.56**	.02
42	.09	**.52**	.11
48	-.06	**.47**	.05
37	.26	**.46**	.02
39	.14	**.45**	.24
12	.20	**.45**	.00
18	.26	**.42**	-.20
15	.06	**.40**	.21
2	.21	.35	-.24
5	.25	.33	.16
13	.16	.31	-.17

Item no	Factors		
	1	2	3
16	.11	.30	.20
52	.06	.28	.20
19	.20	.27	.21
59	.22	.26	.19
22	.15	.19	.02
7	.18	.03	**.56**
34	.25	.15	**.53**
61	.37	.15	**.51**
3	.29	.14	**.49**
43	.17	.01	**.45**
53	.19	.23	**.43**
9	.24	.35	**.43**
29	.20	.01	**.41**
1	-.11	.03	**.40**
54	.29	.14	.34
25	.25	.12	.33
32	.11	-.02	.31
31	.27	.23	.29
47	.23	.22	.25

Note. Loadings ≥. 40 are in boldface.

Appendix B8

Item pool used for confirmatory factor analysis (step 3 of scale development).

Scale 1		Scale 2		Scale 3	
Item	Content	Item	Content	Item	Content
1	Mehrdeutige Situationen machen mich ein wenig nervös.	1	Gute Literatur sollte wenig Raum für Interpretationen zulassen.	1	Die Äußerungen eines Vorgesetzten sollten immer klar und eindeutig sein.
2 (R)	Mehrdeutige Situationen kann ich gut ertragen.	2	Kunst, die viele Interpretationen zulässt, meide ich.	2	Ich beschäftige mich gerne mit eindeutigen Sachverhalten.
3	Ich versuche mehrdeutigen Situationen aus dem Weg zu gehen.	3 (R)	Gute Kunst muss viel Raum für Interpretationen lassen.	3	Ein guter Job ist einer, bei dem immer klar ist, wo was zu tun ist und wie es zu tun ist (Budner, 1962)
4	Ich mag mehrdeutige Situationen nicht. (McLain, 2009)	4 (R)	Ich finde es interessant, über verschiedene potentielle Bedeutungen eines Kunstwerks nachzudenken.	4	Wenn ich eine Situation auf zwei Arten interpretieren kann, habe ich ein großes Bedürfnis, Klarheit herzustellen.
5	Soziale Situationen, die Widersprüche von Mimik, Gestik, Sprache oder Handlung aufweisen, meide ich.	5 (R)	Kunst muss Widersprüche provozieren.	5	Eindeutige Situationen beruhigen mich.
6 (R)	Ich mag es, wenn eine Aussage mehrere Bedeutungen haben kann.	6	Musik, die man auf mehrere Arten interpretieren kann, mag ich nicht.	6 (R)	Soziale Situationen, die nicht eindeutig sind, finde ich spannend.
7	Wenn ich nicht zwischen Ironie und Ernst unterscheiden kann, stresst mich das.	7 (R)	Wenn in einem Roman nicht klar ist, ob das Geschilderte tatsächlich geschehen ist oder ob es nur eine Vorstellung oder ein Traum war, finde ich das spannend.	7 (R)	Eindeutige Sachverhalte finde ich langweilig.
8 (R)	Chatnachrichten oder E-Mails, die Informationen „zwischen den Zeilen" enthalten, finde ich interessant.	8	Wenn eine Zeichnung keine konkreten Gegenstände darstellt, schaue ich sie mir nicht weiter an.	8 (R)	Personen, die versuchen, sich immer klar und eindeutig auszudrücken, finde ich uninteressant.
9	Ich versuche soziale Situationen, in denen Missverständnisse auftreten können, zu vermeiden.	9	Gedichte lese ich ungern.	9	Eindeutigkeit in der Sprache ist immer anzustreben.

Note. (R) Reversed coded item.

Appendix B9

Items per scale after model fit improvement by confirmatory factor analysis (step 3 of scale development)

Scale 1			Scale 2			Scale 3		
Item No	Content	Factor loading	Item No	Content	Factor loading	Item No	Content	Factor loading
1	Mehrdeutige Situationen machen mich ein wenig nervös.	.59	2	Kunst, die viele Interpretationen zulässt, meide ich.	.99	2	Ich beschäftige mich gerne mit eindeutigen Sachverhalten.	.58
2 (R)	Mehrdeutige Situationen kann ich gut ertragen.	.70	3 (R)	Gute Kunst muss viel Raum für Interpretationen lassen.	.60	3	Ein guter Job ist einer, bei dem immer klar ist, wo was zu tun ist und wie es zu tun ist (Budner, 1962)	.58
3	Ich versuche mehrdeutigen Situationen aus dem Weg zu gehen.	.77	4 (R)	Ich finde es interessant, über verschiedene potentielle Bedeutungen eines Kunstwerks nachzudenken.	1	4	Wenn ich eine Situation auf zwei Arten interpretieren kann, habe ich ein großes Bedürfnis, Klarheit herzustellen.	.63
4	Ich mag mehrdeutige Situationen nicht. (McLain, 2009)	.85	8	Wenn eine Zeichnung keine konkreten Gegenstände darstellt, schaue ich sie mir nicht weiter an.	.83	5	Eindeutige Situationen beruhigen mich.	.66
5	Soziale Situationen, die Widersprüche von Mimik, Gestik, Sprache oder Handlung aufweisen, meide ich.	.66						
6 (R)	Ich mag es, wenn eine Aussage mehrere Bedeutungen haben kann.	.59						

Note. (R) Reversed coded item.

Appendix B10

Ambiguous and unambiguous sentences used in the sentence rating task of study 2

Type	Sentence
A	Solch eine gelbe Schale hatte er noch nie gesehen.
U	Solch ein Ruder hatte er noch nie gesehen.
A	Der Hahn machte seltsame Geräusche.
U	Der Stahl machte seltsame Geräusche.
A	Er sagte: „Weizen braucht der Mensch“.
U	Er sagte: „Seide braucht der Mensch“.
A	Diesen Geschmack fand er absonderlich.
U	Dieses Talent fand er absonderlich.
A	Er sah die Maus und fragte sich, wie sie auf den Tisch kam.
U	Er sah die Treppe und fragte sich, wie sie gebaut wurde.
A	Der Mann ging in Richtung Schalter, stoppte dann jedoch.
U	Der Mann ging in Richtung Akte, stoppte dann jedoch.
A	Er ist ein regelmäßiger Besucher der Messe und kennt den Ablauf und den Ort sehr genau.
U	Er kennt den Ablauf und Ort seines Berufes sehr genau.
A	Einen Strauß wollte Sie schon immer bekommen.
U	Ein Zelt wollte sie schon immer bekommen.
A	Ist die Erde der Ursprung des Lebens?
U	Ist das Licht der Ursprung des Lebens?
A	Er sah die Knete an.
U	Er sah den Masseur an.

Note. A = ambiguous. U = Unambiguous.

Appendix B11

Ambiguous sentence beginning and corresponding disambiguating or ambiguous sentences endings used in the sentence choosing task of study 2

Ambiguous sentence	Sentence ending
Solch eine gelbe Schale hatte er noch nie gesehen	und fragte den Verkäufer, was es für eine Frucht sei.
	und fragte den Verkäufer, wer der Hersteller sei.
	und fragte den Verkäufer, wie hoch der Preis sei.
Der Hahn machte seltsame Geräusche,	deshalb ging Paul in den Stall.
	deshalb holte Paul sein Werkzeug.
	deshalb war Paul besorgt.
Er sagte: „Weizen braucht der Mensch“	nach getaner Arbeit.
	für viele Backwaren.
	zum Leben.
Diesen Geschmack fand er absonderlich	da die Gewürze nicht zusammenpassten.
	da die Farben nicht zusammenpassten.
	da er die Kombination so noch nicht kannte.
Er sah die Maus und fragte sich, wie sie auf den Tisch kam,	da er seit langem mit dem Touchpad arbeitete.
	da erst vor kurzem der Kammerjäger da war.
	da er sie lange nicht gesehen hatte.
Der Mann ging in Richtung Schalter, stoppte dann jedoch,	als er bemerkte, dass dieser bereits geschlossen war.
	als er bemerkte, dass dieser bereits gedrückt war.
	als er bemerkte, dass sein Schnürsenkel offen war.
Er ist ein regelmäßiger Besucher der Messe und kennt den Ablauf und den Ort sehr genau,	jedoch hält er von dem Priester nicht viel.
	jedoch hält er von dem Veranstalter nicht viel.
	jedoch sucht er sie nun nicht mehr so regelmäßig auf.
Einen Strauß wollte Sie schon immer bekommen,	da sie Pflanzen liebt.
	da sie Tiere liebt.
	da sie Geschenke liebt.
Ist die Erde der Ursprung des Lebens?	oder doch die Gewässer?
	oder doch ein anderer Planet?
	oder ist das Leben an verschieden Orten entstanden?
Er sah die Knete an	und erinnerte sich an sein verlorenes Geld.
	und erinnerte sich an seine verlorene Kindheit.
	und war gedankenversunken.

Note. The third sentence ending does not disambiguate the initial sentence. The first two sentence endings highlight one of the two meanings of the lexical ambiguous word.

Appendix B12

Overview of items after model fit improvement during confirmatory factor analysis of step 3 (old items) and new added items in step 5 of process of scale development.

Scale 1		Scale 2		Scale 3	
Item	Content	Item	Content	Item	Content
1 Old	**Mehrdeutige Situationen machen mich ein wenig nervös.**	1 Old	**Kunst, die viele Interpretationen zulässt, meide ich.**	1 Old	**Ich beschäftige mich gerne mit eindeutigen Sachverhalten.**
2 (R) Old	**Mehrdeutige Situationen kann ich gut ertragen.**	2 (R) Old	**Gute Kunst muss viel Raum für Interpretationen lassen.**	2 Old	**Ein guter Job ist einer, bei dem immer klar ist, wo was zu tun ist und wie es zu tun ist (Budner, 1962)**
3 Old	**Ich versuche mehrdeutigen Situationen aus dem Weg zu gehen.**	3 (R) Old	**Ich finde es interessant, über verschiedene potentielle Bedeutungen eines Kunstwerks nachzudenken.**	3 Old	**Wenn ich eine Situation auf zwei Arten interpretieren kann, habe ich ein großes Bedürfnis, Klarheit herzustellen.**
4 Old	**Ich mag mehrdeutige Situationen nicht. (McLain, 2009)**	4 Old	**Wenn eine Zeichnung keine konkreten Gegenstände darstellt, schaue ich sie mir nicht weiter an.**	4 Old	**Eindeutige Situationen beruhigen mich.**
5 Old	**Soziale Situationen, die Widersprüche von Mimik, Gestik, Sprache oder Handlung aufweisen, meide ich.**	5 New	Mehrdeutige Kunstwerke bereiten mir weniger Freude als eindeutige.	5 (R) New	Eindeutige Situationen finde ich langweilig.
6 (R) Old	Ich mag es, wenn eine Aussage mehrere Bedeutungen haben kann.	6 New	**Abstrakte Kunst mag ich nicht.**	6 New	**Ich werde unruhig, wenn eine Aufgabe keine eindeutige Lösung hat.**
7 New	Ich mag es nicht, wenn eine Aufgabe mehrere Lösungen zulässt.	7 (R) New	Ich fühle mich wohl, wenn ich mit Kunst in Kontakt komme, die viel Raum für Interpretationen lässt.	7 (R) New	Mir ist es nicht wichtig, eindeutige Strukturen in meinem Leben zu haben.

Note. (R) Reversed coded item. Bold printed items are the final solution with 5 items per scale.

Appendix B13

List of abstract and representational paintings used in validation study (step 6 of scale development)

Artist	Year	Title
	Abstract paintings	
Fer Hakkaart	1969	*Kermistruc*
Fernand Léger	1954	*Two Women Holding Flowers*
Max Ernst	1923	*Men shall know nothing of this*
Tony Tuckson	1970-1973	*White lines (vertical) on ultramarine*
Patrick Caulfield	1969	*Pottery*
Wassily Kandinsky	1913	*Composition VII*
Marc Chagall	1913	*Paris through the Window*
Piet Mondriaan	1921	*Composition en rouge, jaune, bleu et noir*
Giacomo Balla	1912	*Dynamism of a Dog on a Leash, oil on canvas*
Joseph Stella	1913-1914	*Battle of Lights, Coney Island, Mardi Gras*
Pablo Picasso	1938	*Maya with boat*
	Representational paintings	
Lorenzo Costa	1485-1495	*A Concert*
Hendrick Avercamp	1615	*A Scene on the Ice near a Town*
Gerard Houckgeest	1638	*Architectural Fantasy with Figures*
Pieter de Hooch	1663-1665	*Company in a courtyard behind a house*
Gerard ter Borch	1648	*Helena van der Schalcke as a Child*
Limbourg Browsers	-	*Les Tres Riches Heures du duc de Berry (Avril)*
Stubbs	1762	*Mares and foals in a wooded landscape*
Giovanni Battista Tiepolo	1743-1744	*The banquet of Cleopatra*
Eduard Manet	1860	*The ships deck*
Pierre-Auguste Renoir	1881-1886	*The Umbrellas*
Rembrandt	1628	*Two old men disputing*

Appendix C

Comparison of ambiguous and unambiguous words used in Study 1

Word	Freq	No. Letters	Valence Word	Valence (1+2)	Valence 1	Valence 2	Arousal Word	Arousal (1,2)	Arousal 1	Arousal 2	Abst. Word	Abst. (1+2)	Abst. 1	Abst. 2
Anbau	11	5	5.76	5.68	5.26	6.11	2.48	2.65	2.29	3.00	4.14	3.07	2.62	3.52
Kiefer	12	6	5.24	5.63	5.00	6.26	3.56	3.14	2.88	3.39	2.07	1.98	2.19	1.76
Knete	16	5	5.69	6.19	6.09	6.30	3.41	3.71	3.00	4.42	2.00	2.79	2.00	3.57
Strauß	12	6	5.86	6.28	6.71	5.85	3.07	3.74	4.26	3.21	1.90	1.45	1.48	1.43
Viertel	9	7	4.97	5.35	5.51	5.19	2.63	2.64	3.03	2.24	4.76	4.55	3.52	5.57
M	**12.00**	**5.80**	**5.50**	**5.83**	**5.71**	**5.94**	**3.03**	**3.17**	**3.09**	**3.25**	**2.97**	**2.77**	**2.36**	**3.17**
SD	**2.55**	**0.84**	**0.38**	**0.40**	**0.69**	**0.46**	**0.47**	**0.54**	**0.72**	**0.79**	**1.37**	**1.19**	**0.77**	**1.66**
Stahl	11	5	4.78				3.08				1.68			
Zelt	11	4	5.97				3.10				1.35			
Verkauf	9	7	5.06				2.67				4.77			
Masseur	16	7	6.56				3.51				1.48			
Engel	11	5	7.00				2.72				4.32			
M	**11.60**	**5.60**	**5.88**				**3.02**				**2.72**			
SD	**2.61**	**1.34**	**0.95**				**0.34**				**1.68**			
Diff.	**0.40**	**0.20**	**0.37**	**0.05**			**0.01**	**0.16**			**0.25**	**0.04**		

Note. Freq: Frequency; Abst.: Abstractness; Diff.: Difference between mean (ratings) of ambiguous and unambiguous items; Valence (1,2): Mean valence of Valence 1 and Valence 2; Valence 1: Mean valence rating of the fist meaning; Valence 2: Mean valence rating of the second meaning; analogical for Arousal and Abstractness; Meanings of the ambiguous words: “Anbau”: (1) “cultivation farming” (2) “extension to existing buildings”, “Kiefer” (1) “jaw” (2) “pine”, “Knete” (1) “modeling clay” (2) “money”, “Strauß” (1) “bouquet” (2) “ostrich”, “Viertel” (1) “neighborhood” (2) “1/4”; Translation of non-ambiguous words: “Stahl” [“steel”], “Zelt” [“tent”], “Verkauf” [“sale”], “Masseur” [“masseur”], “Engel” [“angel”]

Appendix D1

Sentences used in the learning task in study 2

Matched Target Word		Corresponding two sentences to the target words
Ambiguous	Unambiguous	
Schale	Ruder	Peter holt die/das [Target] aus dem Keller. Langsam legt Tom die/das [Target] beiseite und isst dann seinen Apfel.
Hahn	Stahl	Der [Target] wird mit viel Kraft wieder geradegebogen. Der [Target] ist in der Scheune.
Weizen	Seide	Der/Die [Target] weht im Wind Ein/- [Target] gehört für sie zum Leben dazu.
Geschmack	Talent	In der Kunst hat Sabine kein guten/gutes [Target]. Beim Kochen hilft Thomas sein guter/gutes [Target].
Maus	Treppe	Die [Target] entspricht nicht mehr dem Stand der Technik. Die [Target] quietscht nachts besonders laut.
Schalter	Akte	Die Beamtin stand hinter dem/der [Target]. Im Dunkeln fand Sarah den/die [Target] nicht.
Messe	Beruf	Durch die/den [Target] hat Herr Müller viele innovative Produkte kennengelernt. Laura ist überzeugt, dass die/der [Target] ihr geholfen hat, wieder an das Gute im Menschen zu glauben.
Strauß	Zelt	Zum Geburtstag erhielt Hannelore einen/ein [Target]. Im Nationalpark sah der Ranger in der Ferne einen/ein [Target].
Erde	Licht	Die/Das [Target] legt im Weltraum in kurzer Zeit große Distanzen zurück. In der/dem [Target] gedeiht fast jede Pflanze, sagte der Gärtner.
Knete	Masseur	-/Einen [Target] zu haben ist für manche Menschen sehr wichtig. Im Kindergarten arbeiten/arbeitet sie/- heute mit/ein [Target].

Note. Differences between sentences for the ambiguous and unambiguous target words are marked by "/". A missing word in one version is marked by "-".

Appendix D2

Comparison of ambiguous and unambiguous words used in sequential priming paradigms in Study 2

Word	Freq	No. Letters	Valence Word	Valece (1+2)	Valence 1	Valence 2	Arousal Word	Arousal (1+2)	Arousal 1	Arousal 2	Abs Word	Abs (1+2)	Abs 1	Abs 2
Messe	10	5	5.31	4.94	4.26	5.63	3.30	3.55	3.62	3.48	5.17	4.90	5.71	4.10
Schale	12	6	5.14	5.17	4.89	5.44	2.22	2.42	2.38	2.45	1.72	1.90	1.95	1.86
Schalter	12	8	4.93	5.35	5.37	5.33	3.07	3.05	3.00	3.09	1.86	2.24	1.90	2.57
Hahn	11	4	4.83	5.69	5.74	5.63	2.74	2.77	2.85	2.70	1.83	1.74	1.43	2.05
Maus	12	4	5.28	5.80	5.94	5.67	3.15	2.76	2.85	2.67	1.62	1.64	1.62	1.67
Weizen	13	6	5.86	5.81	6.17	5.44	2.70	2.93	2.32	3.55	1.79	1.76	1.86	1.67
Knete	16	5	5.69	6.19	6.09	6.30	3.41	3.71	3.00	4.42	2.00	2.79	2.00	3.57
Strauß	12	6	5.86	6.28	6.71	5.85	3.07	3.74	4.26	3.21	1.90	1.45	1.48	1.43
Erde	9	4	6.41	6.85	6.51	7.19	3.96	3.94	3.15	4.73	2.17	2.07	1.86	2.29
Geschmack	10	9	6.62	6.86	6.94	6.78	4.63	4.70	4.94	4.45	6.10	5.98	5.14	6.81
M	**11.70**	**5.70**	**5.59**	**5.89**	**5.86**	**5.93**	**3.23**	**3.36**	**3.24**	**3.48**	**2.62**	**2.65**	**2.50**	**2.80**
SD	**1.95**	**1.70**	**0.61**	**0.66**	**0.83**	**0.62**	**0.68**	**0.69**	**0.82**	**0.81**	**1.61**	**1.54**	**1.56**	**1.65**
Beruf	9	5	5.88				3.67				5.1			
Ruder	12	5	5.31				2.18				1.61			
Akte	13	4	4.53				2.77				1.65			
Stahl	11	5	4.78				3.08				1.68			
Treppe	12	6	5.00				2.64				1.42			
Seide	14	5	6.06				2.69				1.71			
Masseur	16	7	6.56				3.51				1.48			
Zelt	11	4	5.97				3.1				1.35			
Licht	8	5	7.31				3.95				2.42			
Talent	11	6	6.88				4.74				6.29			
M	**11.70**	**5.20**	**5.83**				**3.23**				**2.47**			
SD	**2.31**	**0.92**	**0.92**				**0.75**				**1.75**			
Difference	**0.00**	**0.50**	**-0.23**	**0.06**			**-0.01**	**0.13**			**0.15**	**0.18**		

Note. Freq: Frequency; Abs.: Abstractness; Valence (1+2): Mean valence of Valence 1 and Valence 2; Valence 1: Mean valence rating of the fist meaning; Valence 2: Mean valence rating of the second meaning; analogical for Arousal and Abstractness; Meanings of the ambiguous words: “Messe” (1) “mass” (2) “fair”, “Schale” (1) “peel” (2) “bowl”, “Schalter” (1) “switch” (2) “counter window”, “Hahn” (1) “cock (animal)” (2)”cock (technical meaning)”, “Maus” (1) “Mouse (animal) (2) “mouse (computer)”, “Weizen” (1) “wheat” (2) “weiss bier”, “Knete” (1) “modeling clay” (2) “money”, “Strauß” (1) “bouquet” (2) “ostrich”, “Erde” (1) “soil” (2) “earth”, “Geschmack” (1) “taste” (2) “taste (subjective judgment)”; Translation of non-ambiguous words: “Beruf” [“job”], “Ruder” [“oar”], “Akte” [“file”], “Stahl” [“steel”], “Treppe” [“stair”], “Seide” [“silk”], “Masseur” [“masseur”], “Zelt” [“tent”], “Licht” [“light”], “Talent” [“talent”]

Appendix D3

Comparison of negative and positive words used in EP paradigm in Study 2

Word	Freq	No. Letters	Arousal	Abstractness	Valence
Schulden	9	8	5.82	5.03	1.88
Leiche	11	6	5.87	1.65	2.03
Falle	10	5	5.23	3.77	2.22
Knast	12	5	4.92	2.45	2.25
Gift	12	4	5.79	3.13	2.41
Grab	11	4	5.03	1.52	2.59
Ärger	10	5	6.18	5.87	2.69
Angriff	9	7	6.36	4.61	2.69
Tod	8	3	6.31	4.03	2.72
Krise	8	5	5.62	6.16	2.75
Mean	**10.00**	**5.20**	**5.71**	**3.82**	**2.42**
Held	11	4	4.56	5.19	6.88
Sex	10	3	6.28	2.81	7.38
Sommer	8	6	4.44	2.97	7.66
Genuss	11	6	4.87	6.13	7.66
Wunder	9	6	5.13	6.97	7.75
Wärme	11	5	4.59	3.71	7.94
Kuss	13	4	6.26	1.77	7.97
Lächeln	11	7	4.72	2.45	8.25
Freund	9	6	5.18	3.48	8.31
Liebe	8	5	6.33	5.23	8.44
Mean	**10.10**	**5.20**	**5.24**	**4.07**	**7.82**
Difference	**0.10**	**0.00**	**-0.48**	**0.25**	**5.40**

Note. Freq. Frequency; No. Letters: Number of letters; Translation of German words: “Schulden” [“debit”], “Leiche” [“dead body”], “Falle” [“trap”], “Knast” [“clink”], “Gift” [“poison”], “Grab” [“grave”], “Ärger” [“truble”], “Angriff” [“attack”], “Tod” [“death”], “Krise” [“crisis”], “Held” [“hero”], “Sex” [“sex”], “Sommer” [“summer”], “Genuss” [“pleasure”], “Wunder” [“miracle”], “Wärme” [“warmness”], “Kuss” [“kiss”], “Lächeln” [“smile”], “Freund” [“friend”], “Liebe” [“love”]

Appendix D4

Comparison of high and low arousal words used in sequential priming paradigm in Study 2

Word	Freq	No. Letters	Valence	Abstractness	Arousal
Spritze	14	7	3.66	1.39	5.03
Bordell	14	7	3.16	2.61	5.13
Wunder	9	6	7.75	6.97	5.13
Flirt	14	5	6.41	5.65	5.23
Drogen	10	6	3.38	2.94	5.33
Gefahr	8	6	2.97	5.52	6.18
Kuss	13	4	7.97	1.77	6.26
Liebe	8	5	8.44	5.23	6.33
Angriff	9	7	2.69	4.61	6.36
Schrei	13	6	3.25	3.81	6.44
M	**11.20**	**5.90**	**4.97**	**4.05**	**5.74**
SD	**2.62**	**0.99**	**2.37**	**1.84**	**0.61**
Schnur	14	6	4.94	1.61	1.95
Hälfte	8	6	5.03	4.61	2.36
Gebiet	9	6	5.09	5.03	2.38
Zustand	9	7	4.84	7.00	2.38
Taxi	12	4	5.00	1.42	2.44
Anfahrt	13	7	4.53	5.06	2.44
Notiz	14	5	4.97	2.19	2.69
Straße	9	7	5.09	1.65	2.72
Engel	11	5	7.00	4.32	2.72
Nutzen	10	6	5.53	6.52	3.03
M	**10.90**	**5.90**	**5.20**	**3.94**	**2.51**
SD	**2.23**	**0.99**	**0.68**	**2.09**	**0.29**
Difference	**0.30**	**0.00**	**-0.23**	**0.11**	**3.23**

Note. Freq: Frequency; No. Letters: Number of letters; Translation of German words: “Spritze” [“injection”], “Bordell” [“brothel”], “Wunder” [“miracle”], “Flirt” [“flirtation”], “Drogen” [“drugs”], “Gefahr” [“danger”], “Kuss” [“kiss”], “Liebe” [“love”], “Angriff” [“attack”], “Schrei” [“cry”], ”Schnur” [“string”], “Hälfte” [“half”], “Gebiet” [“area”], “Zustand” [“condition”], “Taxi” [“taxi”], “Anfahrt” [“drive”], “Notiz” [“note”], “Straße” [“street”], “Engel” [“angel”], “Nutzen” [“benefit”]

Appendix D5

Comparison of ambiguous and unambiguous words used in IAT in Study 2

Word	Freq	No. Letters	Valence Word	Valence (1+2)	Valence 1	Valence 2	Arousal Word	Arousal (1+2)	Arousal 1	Arousal 2	Abs Word	Abs (1+2)	Abs 1	Abs 2
Schale	12	6	5.14	5.17	4.89	5.44	2.22	2.42	2.38	2.45	1.72	1.90	1.95	1.86
Schalter	12	8	4.93	5.35	5.37	5.33	3.07	3.05	3.00	3.09	1.86	2.24	1.90	2.57
Hahn	11	4	4.83	5.69	5.74	5.63	2.74	2.77	2.85	2.70	1.83	1.74	1.43	2.05
Maus	12	4	5.28	5.80	5.94	5.67	3.15	2.76	2.85	2.67	1.62	1.64	1.62	1.67
Strauß	12	6	5.86	6.28	6.71	5.85	3.07	3.74	4.26	3.21	1.90	1.45	1.48	1.43
M	**11.80**	**5.60**	**5.21**	**5.66**	**5.73**	**5.59**	**2.85**	**2.95**	**3.07**	**2.82**	**1.79**	**1.80**	**1.68**	**1.91**
SD	**0.45**	**1.67**	**0.41**	**0.43**	**0.68**	**0.20**	**0.39**	**0.49**	**0.71**	**0.32**	**0.11**	**0.30**	**0.24**	**0.43**
Ruder	12	5	5.31				2.18				1.61			
Akte	13	4	4.53				2.77				1.65			
Stahl	11	5	4.78				3.08				1.68			
Treppe	12	6	5.00				2.64				1.42			
Zelt	11	4	5.97				3.1				1.35			
M	**11.80**	**4.80**	**5.12**				**2.75**				**1.54**			
SD	**0.84**	**0.84**	**0.56**				**0.38**				**0.15**			
Difference	**0.00**	**0.80**	**0.09**	**0.54**			**0.10**	**0.19**			**0.24**	**0.25**		

Note. Freq: Frequency; No. Letters: Number of letters; Abs: Abstractness; Valence (1+2): Mean valence of Valence 1 and Valence 2; Valence 1: Mean valence rating of the first meaning; Valence 2: Mean valence rating of the second meaning; analogical for arousal and abstractness; Meanings of the ambiguous words: “Schale” (1) “peel” (2) “bowl”, “Schalter” (1) “switch” (2) “counter window”, “Hahn” (1) “cock (animal)” (2)”cock (technical meaning)”, “Maus” (1) “Mouse (animal) (2) “mouse (computer)”, “Strauß” (1) “bouquet” (2) “ostrich”; Translation of non-ambiguous words: “Ruder” [“oar”], “Akte” [“file”], “Stahl” [“steel”], “Treppe” [“stair”], , “Zelt” [“tent”]

Appendix D6

Comparison of negative and positive words used in IAT in Study 2

Word	Freq	No. Letters	Arousal	Abstractness	Valence
Tumor	13	5	6.05	2.52	1.50
Falle	10	5	5.23	3.77	2.22
Knast	12	5	4.92	2.45	2.25
Ärger	10	5	6.18	5.87	2.69
Krise	8	5	5.62	6.16	2.75
M	**10.60**	**5.00**	**5.60**	**4.15**	**2.28**
SD	**1.95**	**0.00**	**0.53**	**1.78**	**0.50**
Wunder	9	6	5.13	6.97	7.75
Wärme	11	5	4.59	3.71	7.94
Kuss	13	4	6.26	1.77	7.97
Lächeln	11	7	4.72	2.45	8.25
Liebe	8	5	6.33	5.23	8.44
M	**10.40**	**5.40**	**5.41**	**4.03**	**8.07**
SD	**1.95**	**1.14**	**0.84**	**2.11**	**0.27**
Difference	**-0.20**	**0.40**	**-0.19**	**-0.13**	**5.79**

Note. Freq: Frequency; No. Letters: Number of letters; Translation of German words: “Tumor” [“tumor”], “Falle” [“trap”], “Knast” [“clink”], “Ärger” [“truble”], “Krise” [“crisis”], “Wunder” [“miracle”], “Wärme” [“warmness”], “Kuss” [“kiss”], “Lächeln” [“smile”], “Liebe” [“love”]

Appendix D7

Comparison of high and low arousal words used in the IAT in Study 2

Word	Freq	No. Letters	Valence	Abstractness	Arousal
Spritze	14	7	3.66	1.39	5.03
Wunder	9	6	7.75	6.97	5.13
Flirt	14	5	6.41	5.65	5.23
Drogen	10	6	3.38	2.94	5.33
Gefahr	8	6	2.97	5.52	6.18
M	**11.00**	**6.00**	**4.83**	**4.49**	**5.38**
SD	**2.83**	**0.71**	**2.12**	**2.27**	**0.46**
Hälfte	8	6	5.03	4.61	2.36
Anfahrt	13	7	4.53	5.06	2.44
Notiz	14	5	4.97	2.19	2.69
Straße	9	7	5.09	1.65	2.72
Nutzen	10	6	5.53	6.52	3.03
M	**10.80**	**6.20**	**5.03**	**4.01**	**2.65**
SD	**2.59**	**0.84**	**0.36**	**2.04**	**0.26**
Difference	**0.20**	**-0.20**	**-0.20**	**0.49**	**2.73**

Note. Freq: Frequency; No. Letters: Number of letters; Translation of German words: “Spritze” [“injection”], “Wunder” [“miracle”], “Flirt” [“flirtation”], “Drogen” [“drugs”], “Gefahr” [“danger”], “Hälfte” [“half”], “Anfahrt” [“drive”], “Notiz” [“note”], “Straße” [“street”], “Nutzen” [“benefit”]

Appendix E1

Comparison of ambiguous and unambiguous words used in sequential priming paradigms in Study 3

Ambiguous Words															Unambiguous Words					
Word	Freq	No. Lett	V (M)	V (W)	V 1	V 2	Aro (M)	Aro (W)	Aro 1	Aro 2	Abs (M)	Abs (W)	Abs 1	Abs 2	Word	Freq	No. Lett	V	Aro	Abs
Wanze	17	5	3.13	3.45	3.14	3.11	4.76	3.78	3.79	5.73	1.90	2.45	1.67	2.14	Bordell	14	7	3.16	5.13	2.61
Steuer	10	6	3.67	3.90	4.46	2.89	3.95	4.30	3.56	4.33	4.02	5.69	2.76	5.29	Drogen	10	6	3.38	5.33	2.94
Korb	12	4	4.07	4.69	5.14	3.00	3.78	2.67	2.29	5.27	4.05	1.59	1.86	6.24	Spritze	14	7	3.66	5.03	1.39
Schlange	12	8	4.13	4.41	4.77	3.48	4.22	4.81	4.56	3.88	2.62	1.90	1.52	3.71	Harem	16	5	4.16	4.41	3.9
Geschoss	14	8	4.22	3.97	3.17	5.26	4.42	5.30	6.24	2.61	2.76	3.38	2.43	3.10	Sau	14	3	4.22	3.54	1.68
Mast	14	4	4.49	4.93	3.80	5.19	3.24	2.93	3.97	2.52	3.57	2.21	4.38	2.76	Anfahrt	13	7	4.53	2.44	5.06
Kluft	13	5	5.02	4.14	4.46	5.59	3.55	4.78	3.97	3.12	2.71	4.59	2.71	2.71	Notiz	14	5	4.97	2.69	2.19
Pflaster	13	8	5.10	4.52	4.69	5.52	2.90	3.74	2.68	3.12	1.98	1.62	2.33	1.62	Treppe	12	6	5.00	2.64	1.42
Schale	12	6	5.17	5.14	4.89	5.44	2.42	2.22	2.38	2.45	1.90	1.72	1.95	1.86	Slalom	12	6	5.03	3.13	3.23
Diele	16	5	5.22	4.86	5.37	5.07	2.37	2.85	2.44	2.30	2.50	2.86	1.76	3.24	Verkauf	9	7	5.06	2.67	4.77
Stock	11	5	5.33	4.86	5.26	5.41	2.64	2.63	2.79	2.48	2.12	1.93	1.38	2.86	Straße	10	6	5.09	2.72	1.65
Viertel	9	7	5.35	4.97	5.51	5.19	2.64	2.63	3.03	2.24	4.55	4.76	3.52	5.57	Ruder	12	5	5.31	2.18	1.61
Radler	12	6	5.63	5.55	5.49	5.78	3.67	3.44	3.71	3.64	1.86	2.55	2.05	1.67	Busen	14	5	5.41	3.92	1.81
Hahn	11	4	5.69	4.83	5.74	5.63	2.77	2.74	2.85	2.70	1.74	1.83	1.43	2.05	Zelt	11	4	5.97	3.1	1.35
Maus	12	4	5.80	5.28	5.94	5.67	2.76	3.15	2.85	2.67	1.64	1.62	1.62	1.67	Brief	10	5	6.03	3.21	1.29
Weizen	13	6	5.81	5.86	6.17	5.44	2.93	2.70	2.32	3.55	1.76	1.79	1.86	1.67	Seide	14	5	6.06	2.69	1.71
Strauß	12	6	6.28	5.86	6.71	5.85	3.74	3.07	4.26	3.21	1.45	1.90	1.48	1.43	Ofen	12	4	6.25	2.74	1.29
Weide	13	5	6.46	6.17	6.40	6.52	2.66	2.89	2.24	3.09	2.29	2.31	2.00	2.57	Hafen	10	5	6.47	2.59	1.87
Erde	9	4	6.85	6.41	6.51	7.19	3.94	3.96	3.15	4.73	2.07	2.17	1.86	2.29	Baby	10	4	6.47	4.13	1.26
Kiwi	16	4	7.06	5.83	7.09	7.04	3.35	3.67	3.12	3.58	1.40	1.79	1.38	1.43	Masseur	16	7	6.56	3.51	1.48
M	**12.55**	**5.50**	**5.22**	**4.98**	**5.24**	**5.21**	**3.34**	**3.41**	**3.31**	**3.36**	**2.45**	**2.53**	**2.10**	**2.79**	***M***	**12.35**	**5.45**	**5.14**	**3.39**	**2.23**
SD	**2.14**	**1.40**	**1.04**	**0.80**	**1.10**	**1.22**	**0.71**	**0.86**	**0.98**	**1.00**	**0.92**	**1.17**	**0.77**	**1.41**	***SD***	**2.11**	**1.19**	**1.04**	**0.96**	**1.17**
Diff	**0.20**	**0.05**	**0.08**	**-0.16**			**-0.05**	**0.02**			**0.22**	**0.31**								

Note. Freq: Frequency; Diff: Difference of mean ratings of ambiguous and unambiguous words; Aro: Arousal; Abs: Abstractness; V(W): Valence of the Word; V(M): Mean of V(M1) and V(M2); V(M1): Mean valence rating of the first meaning; V(M2): Mean valence rating of the second meaning; analogical for arousal and Abstractness; Meanings of the ambiguous words: “Wanze” (1) “bug (animal)” (2) “bugging device”, “Steuer” (1) “steering wheel” (2) “tax”, “Korb” (1) “basked” (2) “to get the brush-off”, “Schlange” (1) “snake” (2) “queue”, “Geschoss” (1) “bullet” (2) “floor”, “Mast” (1) “mast [AGR.]” (2) “pole”, “Kluft” (1) “gap” (2) “cleavage [fig.]”, “Pflaster” (1) “road coating” (2) “plaster”, “Schale” (1) “peel” (2) “bowl”, “Diele” (1)”bord” (2) “entry”, “Stock” (1) “stick” (2) “floor”, “Viertel” (1) “neighborhood” (2) “fourth part”, “Radler” (1) “cyclist” (2) “shandy”, “Hahn” (1) “cock (animal)” (2)”cock (technical meaning)”, “Maus” (1) “Mouse (animal) (2) “mouse (computer)”, “Weizen” (1) “wheat” (2) “weiss bier”, “Strauß” (1) “bouquet” (2) “ostrich”, “Weide” (1) “osier” (2) “meadow”, “Erde” (1) “soil” (2) “earth”, “Kiwi” (1) “kiwi (animal)”, (2) “kiwi (fruit)”; Translation of non-ambiguous words: “Bordell” [“brothel”], “Drogen” [“drugs”], “Spritze” [“injection”], “Harem” [“harem”], “Sau” [“sow (animal)”], “Anfahrt” [“drive”], “Notiz” [“note”], “Treppe” [“stair”], “Slalom” [“slalom”], “Verkauf” [“selling”], “Straße” [“street”], “Ruder” [“oar”], “Busen” [“breast”], “Zelt” [“tent”], “Brief” [“letter”], “Seide” [“silk”], “Ofen” [“stove”], “Hafen” [“harbor”], “Baby” [“baby”], “Masseur” [“masseur”]

Appendix E2

Comparison of positive and negative words used in sequential priming paradigms in Study 3

Negative Words						Positive Words					
Word	Freq	Letters	Arousal	Abst	Valence	Word	Freq	Letters	Arousal	Abst	Valence
Schulden	9	8	5.82	5.03	1.88	Ausflug	11	7	3.82	4.39	7.25
Leiche	11	6	5.87	1.65	2.03	Sex	10	3	6.28	2.81	7.38
Knast	12	5	4.92	2.45	2.25	Genuss	11	6	4.87	6.13	7.66
Grab	11	4	5.03	1.52	2.59	Wunder	9	6	5.13	6.97	7.75
Ärger	10	5	6.18	5.87	2.69	Wärme	11	5	4.59	3.71	7.94
Angriff	9	7	6.36	4.61	2.69	Kuss	13	4	6.26	1.77	7.97
Krise	8	5	5.62	6.16	2.75	Lächeln	11	7	4.72	2.45	8.25
Asthma	13	6	4.64	3.77	2.78	Freund	9	6	5.18	3.48	8.31
Geiz	14	4	4.49	7.10	2.91	Liebe	8	5	6.33	5.23	8.44
Gefahr	8	6	6.18	5.52	2.97	Freiheit	9	8	5.49	6.84	8.47
M	**10.50**	**5.60**	**5.51**	**4.37**	**2.55**	***M***	**10.20**	**5.70**	**5.27**	**4.38**	**7.94**
SD	**2.07**	**1.26**	**0.69**	**1.95**	**0.37**	***SD***	**1.48**	**1.49**	**0.83**	**1.85**	**0.43**
Difference	**-0.30**	**0.10**	**-0.24**	**0.01**	**5.39**						

Note. Freq: Frequency; Letters: number of letters; Abst: Abstractness; Difference: Difference of mean ratings of negative and positive words; Translation of German words: “Schulden” [“debit”], “Leiche” [“dead body”], “Knast” [“clink”], “Grab” [“grave”], “Ärger” [“truble”], “Angriff” [“attack”], “Krise” [“crisis”], “Asthma” [“asthma”], “Geiz” [“meanness”], “Gefahr” [“danger”], “Ausflug” [“trip”], “Sex” [“sex”], “Genuss” [“pleasure”], “Wunder” [“miracle”], “Wärme” [“warmness”], “Kuss” [“kiss”], “Lächeln” [“smile”], “Freund” [“friend”], “Liebe” [“love”], “Freiheit” [“freedom”]

Appendix E3

Comparison of ambiguous and unambiguous words used in the IAT in Study 3

Ambiguous Words															Unambiguous Words					
Word	Freq	No. Lett	V (M)	V (W)	V 1	V 2	Aro (M)	Aro (W)	Aro 1	Aro 2	Abs (M)	Abs (W)	Abs 1	Abs 2	Word	Freq	No. Lett	V	Aro	Abs
Kluft	13	5	5.02	4.14	4.46	5.59	3.55	4.78	3.97	3.12	2.71	4.59	2.71	2.71	Notiz	14	5	4.97	2.69	2.19
Schale	12	6	5.17	5.14	4.89	5.44	2.42	2.22	2.38	2.45	1.90	1.72	1.95	1.86	Straße	10	6	5.09	2.72	1.65
Stock	11	5	5.33	4.86	5.26	5.41	2.64	2.63	2.79	2.48	2.12	1.93	1.38	2.86	Ruder	12	5	5.31	2.18	1.61
Radler	12	6	5.63	5.55	5.49	5.78	3.67	3.44	3.71	3.64	1.86	2.55	2.05	1.67	Busen	14	5	5.41	3.92	1.81
Hahn	11	4	5.69	4.83	5.74	5.63	2.77	2.74	2.85	2.70	1.74	1.83	1.43	2.05	Zelt	11	4	5.97	3.1	1.35
M	**11.80**	**5.20**	**5.37**	**4.90**	**5.17**	**5.57**	**3.01**	**3.16**	**3.14**	**2.88**	**2.07**	**2.52**	**1.90**	**2.23**	***M***	**12.20**	**5.00**	**5.35**	**2.92**	**1.72**
SD	**0.84**	**0.29**	**0.52**	**0.51**	**0.15**	**0.56**	**1.00**	**0.67**	**0.50**	**0.39**	**1.20**	**0.54**	**0.53**	**0.84**	***SD***	**1.79**	**0.71**	**0.39**	**0.65**	**0.31**
Diff	**-0.40**	**0.20**	**0.02**	**-0.45**			**0.09**	**0.24**			**0.34**	**0.80**								

Note. Freq: Frequency; No Lett: Number of letters; Aro: Arousal; Abs: Abstractness; V(W): Valence of the Word; V(M): Mean of V1 and V2; V1: Mean valence rating of the first meaning; V2: Mean valence rating of the second meaning; analogical for arousal and abstractness; Meanings of the ambiguous words: “Kluft” (1) “gap” (2) “cleavage [fig.]”, “Schale” (1) “peel” (2) “bowl”, “Stock” (1) “stick” (2) “floor”, “Radler” (1) “cyclist” (2) “shandy”, Hahn” (1) “cock (animal)” (2)”cock (technical meaning)”; Translation of unambiguous words: “Notiz” [“note”], “Straße” [“street”], “Ruder” [“oar”], “Busen” [“breast”], “Zelt” [“tent”]

Appendix E4

Comparison of positive and negative words used in the IAT in Study 3

Negative Words						Positive Words					
Word	Freq	Letters	Arousal	Abs	Valence	Word	Freq	Letters	Arousal	Abs	Valence
Leiche	11	6	5.87	1.65	2.03	Wärme	11	5	4.59	3.71	7.94
Knast	12	5	4.92	2.45	2.25	Kuss	13	4	6.26	1.77	7.97
Ärger	10	5	6.18	5.87	2.69	Lächeln	11	7	4.72	2.45	8.25
Angriff	9	7	6.36	4.61	2.69	Freund	9	6	5.18	3.48	8.31
Asthma	13	6	4.64	3.77	2.78	Liebe	8	5	6.33	5.23	8.44
M	**11.00**	**5.80**	**5.59**	**3.67**	**2.49**	***M***	**10.40**	**5.40**	**5.42**	**3.33**	**8.18**
SD	**1.58**	**0.84**	**0.77**	**1.68**	**0.33**	***SD***	**1.95**	**1.14**	**0.83**	**1.32**	**0.22**
Difference	**-0.60**	**-0.40**	**-0.18**	**-0.34**	**5.69**						

Note. Freq: Frequency; Letters: Number of letters; Abs: Abstractness; Translation of German words: “Leiche” [“dead body”], “Knast” [“clink”], “Grab” [“grave”], “Ärger” [“truble”], “Angriff” [“attack”], “Asthma” [“asthma”], “Wärme” [“warmness”], “Kuss” [“kiss”], “Lächeln” [“smile”], “Freund” [“friend”], “Liebe” [“love”]

Appendix F

Information provided in the pre-registration of study 4 on asPredicted.org

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?

We are investigating whether lexically ambiguous and unambiguous words trigger an automatic activation of valence (as assessed with an evaluative priming [EP] paradigm) depending on the associations of the concept's ambiguity and clearness with positive or negative valence, respectively.

In order to assess the valence of associations with these concepts, we use the Implicit Association Test (IAT). In block 3 and 4, negative + ambiguous and positive + unambiguous will share one response key. In block 6 and 7, negative + unambiguous and positive + ambiguous will share one response key.

For the IAT and EP paradigm, we use ambiguous words matched with unambiguous words for frequency, length, valence, arousal, and abstractness (based on a pretest). These ambiguous and unambiguous words serve as primes (duration: 200 ms; stimulus onset asynchrony: 300 ms) in an EP task and as target concepts in the IAT. Positive and negative words matched for frequency, length, arousal, and abstractness are used as targets in the EP paradigm and as attribute dimensions in the IAT.

The stimulus set for the EP paradigm consists of 20 ambiguous, 20 unambiguous, 10 positive, and 10 negative words. This paradigm has 40 training trials followed by three test blocks (80 trials each). The stimulus set used for the IAT is a subset of the EP stimulus set consisting of 5 ambiguous, 5 unambiguous, 5 positive, and 5 negative words.

Hypotheses:

(1) The Prime x Target interaction is moderated by the IAT D Value:

- If the IAT indicates the association that ambiguity = negative and clearness = positive, we expect the following prime × target interaction: if an ambiguous prime precedes a negative target, the reaction time should be faster than if an unambiguous prime precedes a negative target. If an ambiguous prime precedes a positive target, the reaction time should be slower compared to an unambiguous prime preceding a positive target.

  If the IAT indicates the association that ambiguity = positive and clearness = negative, the prime × target interaction should indicate the reversed pattern.

3) Describe the key dependent variable(s) specifying how they will be measured.

Latencies in the EP paradigm constitute the key dependent variable.

4) How many and which conditions will participants be assigned to?
2 (prime type: ambiguous vs. unambiguous) × 2 (target type: positive vs. negative) design with repeated measures on both factors.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

We will conduct multi-level analyses with random intercepts for the 40 prime words, 20 target words and participants. In a fixed effects model, we will regress the logarithmized reaction times in the EP paradigm on prime type × target type × IAT D-value and all subordinate two-way interactions and the main effects. In order to bind error variance, we will include a two-way interaction of the target type with the target type of the previous trial. The target type of the previous trial as a covariate is also included. In the same vain, we will include a two-way interaction of the valence of the prime with the target type and the valence of the prime as a covariate.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

Data of participants who did not finish the online study will not be used. All trials in the EP paradigm with reaction times below 300 and above 3000 ms and all false classifications will be excluded from analysis.

7) How many observations will be collected or what will determine sample size?
No need to justify decision, but be precise about exactly how the number will be determined.

A link to the online study will be published via the mailing list of the university (students and employees) and will be active for exactly one week (168 hours) before it will be deactivated. Based on previous studies, we expect to collect data from about 300 participants within this period.

8) Anything else you would like to pre-register?
(e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

In terms of IAT analysis, we will use the improved algorithm (Greenwald, Nosek & Banaji, 2003, Table 4) to calculate the D score. That score will be used in the multi-level analyses (see above).

Appendix G1

Information provided in the pre-registration of study 5 on asPredicted.org

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?

We are investigating whether lexically ambiguous and unambiguous (clear) words trigger an automatic activation of valence (as assessed with an evaluative priming [EP] paradigm) depending on the induced association of the concepts ambiguity and clearness with positive or negative valence, respectively.

For the induction of valence associations with these concepts, we use the first three blocks of the Implicit Association Test. In one condition, negative + unambiguous and positive + ambiguous will share one response key. In the other condition, negative + ambiguous and positive + unambiguous will share one response key.

For the EP paradigm we use ambiguous words matched with unambiguous words for frequency, length, valence, arousal, abstractness (based on a pretest). These ambiguous and unambiguous words serve as primes in an EP task. Positive and negative words matched for frequency, length, arousal, abstractness are used as targets in the EP paradigm.

We use two different stimulus sets (A, B). Each set consists of 6 ambiguous, 6 unambiguous, 6 positive, and 6 negative words. We counterbalance the specific set used for learning across participants. Stimulus sets A and B are mixed within each block of the EP task. There are two blocks (96 trials each). Within each trial prime and target words belong to the same stimulus set. Therefore, half of the stimuli in the EP paradigm were used for learning and half are new stimuli. The known stimuli serve to check whether the EP paradigm is sensitive to the manipulation. The new stimuli serve to investigate whether our induction of associations generalizes to new ambiguous and unambiguous words (primes in the EP paradigm).

Hypotheses:

(2) The Prime x Target interaction is moderated by the induced associations

- In the ambiguity = negative and clearness = positive condition, we expect the following prime × target interaction: if an ambiguous prime precedes a negative target, the reaction time should be faster than if an unambiguous prime precedes a negative target. If an ambiguous prime precedes a positive target, the reaction time should be slower compared to an unambiguous prime preceding a positive target.
- In the ambiguity = positive and clearness = negative condition, the prime × target interaction should indicate the reversed pattern: if an ambiguous prime precedes a negative target, the reaction time should be slower than if an unambiguous prime precedes a negative target. If an ambiguous prime precedes a positive target, the reaction time should be faster compared to an unambiguous prime preceding a positive target.

(3) The Prime × Target × Induced Association interaction is further moderated by the novelty of the stimuli

- For stimuli used in the induction phase, the three-way-interaction postulated in (1) should be stronger than for new stimuli.

3) Describe the key dependent variable(s) specifying how they will be measured.

Latencies in the EP paradigm constitute the key dependent variable.

4) How many and which conditions will participants be assigned to?
2 (Induced Association: ambiguity=positive and clearness=negative vs. ambiguity=negative and clearness=positive) × 2 (prime type: ambiguous vs. unambiguous) × 2 (target type: positive vs. negative) × 2 (stimulus novelty: same as in induction phase vs. different from induction phase) design with repeated measures on the last three factors. Additionally, across participants we counterbalance the specific stimulus set used for learning (see above).

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

We will conduct multi-level analyses with random intercepts for the 24 prime words, 24 target words and participants. In the fixed effects model, we regress the logarithmized reaction times in the EP paradigm on prime type × target type × induction × stimulus novelty × stimulus set used in the induction phase and all subordinate four-way, three-way, two-way interactions and the main effects. In order to bind error variance, we will include a two-way interaction of the target type with the target type of the previous trial. The target type of the previous trial as a main effect is also included. If the stimulus set (A or B) used in the induction phase does not interact with the three-way interaction specified in Hypothesis 1 and with the four-way interaction specified in Hypothesis 2, it will be dropped from the model.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

Data of participants who did not finish the online study will not be used. All trials in the EP paradigm with reaction times below 300 and above 3000ms and all false classifications will be excluded from analysis.

7) How many observations will be collected or what will determine sample size?
No need to justify decision, but be precise about <u>exactly</u> how the number will be determined.

The link to the online study will be published via the mailing list of the university (students and employees) and will be active for exactly one week (168 hours) before it will be deactivated. Based on previous studies, we expect to collect data from about 400 participants within this period.

8) Anything else you would like to pre-register?
(e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

As a manipulation check, we will conduct an Implicit Association Test after the EP paradigm and expect that the induction will have an impact on the D score. We will use the improved algorithm (Greenwald, Nosek & Banaji, 2003, Table 4).

Appendix G2

Comparison of ambiguous and unambiguous words of stimuli set A and B used in study 5

Ambiguous Words														Unambiguous Words						
Word	Freq	No. Lett	V (M)	V (W)	V 1	V 2	Aro (M)	Aro (W)	Aro 1	Aro 2	Abs (M)	Abs (W)	Abs 1	Abs 2	Word	Freq	No. Lett	V	Aro	Abs
Stimuli set A																				
Bank	8	4	4.95	4.79	3.86	6.04	3.12	3.93	3.79	2.45	2.83	2.28	3.71	1.95	Stahl	11	5	4.78	3.08	1.68
Sender	10	6	4.98	4.90	4.86	5.11	3.22	3.11	3.38	3.06	4.21	4.45	4.24	4.19	Notiz	14	5	4.97	2.69	2.19
Pflaster	13	8	5.10	4.52	4.69	5.52	2.90	3.74	2.68	3.12	1.98	1.62	2.33	1.62	Hälfte	8	6	5.03	2.36	4.61
Orden	12	5	5.20	5.86	4.40	6.00	3.72	3.74	3.50	3.94	4.43	3.17	5.52	3.33	Verkauf	9	7	5.06	2.67	4.77
Zylinder	14	8	5.25	5.52	4.94	5.56	2.73	3.56	2.94	2.52	1.88	1.97	1.95	1.81	Spiegel	10	7	5.25	3.15	1.26
Grund	7	5	5.47	5.17	5.09	5.85	3.10	2.85	2.50	3.70	4.50	6.14	2.71	6.29	Butter	11	6	5.28	2.05	1.26
M	**10.67**	**6.00**	**5.16**	**5.13**	**4.64**	**5.68**	**3.13**	**3.49**	**3.13**	**3.13**	**3.31**	**3.27**	**3.41**	**3.20**	***M***	**10.50**	**6.00**	**5.06**	**2.67**	**2.63**
SD	**2.80**	**1.67**	**0.19**	**0.50**	**0.45**	**0.35**	**0.34**	**0.42**	**0.51**	**0.60**	**1.23**	**1.73**	**1.34**	**1.82**	***SD***	**2.07**	**0.89**	**0.19**	**0.42**	**1.63**
Diff	**0.17**	**0.00**	**0.10**	**0.06**			**0.47**	**0.82**			**0.68**	**0.64**								
Stimuli set B																				
Verband	10	7	4.97	4.45	4.60	5.33	3.77	3.70	3.94	3.61	4.02	2.55	2.48	5.57	Slalom	12	6	5.03	3.13	3.23
Kluft	13	5	5.02	4.14	4.46	5.59	3.55	4.78	3.97	3.12	2.71	4.59	2.71	2.71	Korken	15	6	5.06	2.41	1.19
Schale	12	6	5.17	5.14	4.89	5.44	2.42	2.22	2.38	2.45	1.90	1.72	1.95	1.86	Straße	10	6	5.09	2.72	1.65
Diele	16	5	5.22	4.86	5.37	5.07	2.37	2.85	2.44	2.30	2.50	2.86	1.76	3.24	Gebiet	9	6	5.09	2.38	5.03
Schalter	12	8	5.35	4.93	5.37	5.33	3.05	3.07	3.00	3.09	2.24	1.86	1.90	2.57	Ruder	12	5	5.31	2.18	1.61
Kapelle	11	7	5.47	5.28	5.09	5.85	3.30	3.11	2.91	3.70	2.26	2.14	1.67	2.86	Busen	14	5	5.41	3.92	1.81
M	**12.33**	**6.33**	**5.20**	**4.80**	**4.96**	**5.44**	**3.08**	**3.29**	**3.11**	**3.05**	**2.61**	**2.62**	**2.08**	**3.13**	***M***	**12.00**	**5.67**	**5.17**	**2.79**	**2.42**
SD	**2.07**	**1.21**	**0.19**	**0.43**	**0.39**	**0.26**	**0.58**	**0.87**	**0.70**	**0.57**	**0.75**	**1.05**	**0.42**	**1.28**	***SD***	**2.28**	**0.52**	**0.16**	**0.65**	**1.46**
Diff	**0.33**	**0.67**	**0.04**	**-0.37**			**0.29**	**0.50**			**0.19**	**0.20**								

Note. Freq: Frequency; No. Lett.: Number of letters; Aro: Arousal, Abs: Abstractness; V(W): Valence of the Word; V(M): Mean of V1 and V2; V1: Mean valence rating of the first meaning; V2: Mean valence rating of the second meaning; for arousal and abstractness analogical; Meanings of the ambiguous words: “Bank” (1) “bank [FINAN.]” (2) “bench”, “Sender” (1) “transmitter [TECH.]”, “ broadcast station [TELEKOM.]”, “Pflaster” (1) “road coating” (2) “plaster”, “Orden” (1) “fraternity” (2) “decoration [e.g. MILIT.]”, “Zylinder” (1) “cylinder [TECH.]” (2) “top hat”, “Grund” (1) “ground” (2) “reason”, “Verband” (1) “bandage [MED.]” (2) “federation”, Kluft” (1) “gap” (2) “cleavage [fig.]”, “Schale” (1) “peel” (2) “bowl”, “Diele” (1)”bord” (2) “entry”, “Schalter” (1) “switch [ELEKT.]” (2) “counter window”, “Kapelle” (1) “chapel [REL.] (2) “band [MUS.]”; Translation of unambiguous words: “Stahl” [“steel”], “Notiz” [“note”], “Hälfte” [“half”], “Verkauf” [“selling”], “Spiegel” [“mirror”], “Butter” [“butter”], “Slalom” [“slalom”], “Korken” [“cork”], “Straße” [“street”], “Gebiet” [“area”], “Ruder” [“oar”], “Busen” [“breast”],

Appendix G3

Comparison of positive and negative words of stimuli set A and B used in study 5

Negative Words							Positive Words					
Word	Freq	Letters	Arousal	Abst.	Valence		Word	Freq	Letters	Arousal	Abst.	Valence
						Stimuli set A						
Schulden	9	8	5.82	5.03	1.88		Licht	8	5	3.95	2.42	7.31
Leiche	11	6	5.87	1.65	2.03		Sommer	8	6	4.44	2.97	7.66
Henker	15	6	5.33	2.61	2.22		Wunder	9	6	5.13	6.97	7.75
Tod	8	3	6.31	4.03	2.72		Kuss	13	4	6.26	1.77	7.97
Asthma	13	6	4.64	3.77	2.78		Lächeln	11	7	4.72	2.45	8.25
Gefahr	8	6	6.18	5.52	2.97		Liebe	8	5	6.33	5.23	8.44
M	**10.67**	**5.83**	**5.69**	**3.77**	**2.43**		***M***	**9.50**	**5.50**	**5.14**	**3.63**	**7.90**
SD	**2.88**	**1.60**	**0.62**	**1.45**	**0.45**		***SD***	**2.07**	**1.05**	**0.98**	**2.02**	**0.41**
Difference	**1.17**	**0.33**	**0.56**	**0.13**	**-5.46**							
						Stimuli set B						
Falle	10	5	5.23	3.77	2.22		Sex	10	3	6.28	2.81	7.38
Knast	12	5	4.92	2.45	2.25		Genuss	11	6	4.87	6.13	7.66
Grab	11	4	5.03	1.52	2.59		Wärme	11	5	4.59	3.71	7.94
Angriff	9	7	6.36	4.61	2.69		Meer	9	4	3.79	1.48	8.19
Krise	8	5	5.62	6.16	2.75		Freund	9	6	5.18	3.48	8.31
Geiz	14	4	4.49	7.10	2.91		Freiheit	9	8	5.49	6.84	8.47
M	**10.67**	**5.00**	**5.27**	**4.27**	**2.57**		***M***	**9.83**	**5.33**	**5.03**	**4.08**	**7.99**
SD	**2.16**	**1.10**	**0.65**	**2.14**	**0.28**		***SD***	**0.98**	**1.75**	**0.84**	**2.03**	**0.42**
Difference	**-0.83**	**0.33**	**-0.24**	**-0.19**	**5.42**							

Note. Freq: Frequency; Letters: Number of letters; Abst.: Abstractness; Translation of words: “Schulden” [“debit”], “Leiche” [“dead body”], “Henker” [“executioner”], “Tod” [“death”], “Asthma” [“asthma”], “Gefahr” [“danger”], “Falle” [“trap”], “Knast” [“clink”], “Grab” [“grave”], “Angriff” [“attack”], “Krise”, “Geiz” [“meanness”], “Licht” [“light”], “Sommer” [“summer”], “Wunder” [“miracle”], “Kuss” [“kiss”], “Lächeln” [“smile”], “Liebe” [“love”], “Sex” [“sex”], “Genuss” [“pleasure”], “Wärme” [“warmness”], “Meer” [“ocean”], “Freund” [“friend”], “Freiheit” [“freedom”]

Appendix H

Information provided in the pre-registration of study 6 on asPredicted.org

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
We are investigating whether lexically ambiguous and unambiguous (clear) words trigger an automatic activation of valence (as assessed with an evaluative priming [EP] paradigm) depending on the induced association of the concepts ambiguity and clearness with positive or negative valence, respectively.
For the induction of valence associations with these concepts, we use the first three blocks of the Implicit Association Test. In one condition, negative + unambiguous and positive + ambiguous will share one response key. In the other condition, negative + ambiguous and positive + unambiguous will share one response key.
For the EP paradigm we use ambiguous words matched with unambiguous words for frequency, length, valence, arousal, abstractness (based on a pretest). These ambiguous and unambiguous words serve as primes in an EP task. Positive and negative words matched for frequency, length, arousal, abstractness are used as targets in the EP paradigm.
We use two different stimulus sets (A, B). Each set consists of 6 ambiguous, 6 unambiguous, 6 positive, and 6 negative words. We counterbalance the specific set used for learning across participants. Stimulus sets A and B are mixed within each block of the EP task. There are two blocks (96 trials each) with different SOA (200ms vs. 400ms). The sequence of SOA is counterbalanced. Within each trial prime and target words belong to the same stimulus set. Therefore, half of the stimuli in the EP paradigm were used for learning and half are new stimuli. The known stimuli serve to check whether the EP paradigm is sensitive to the manipulation. The new stimuli serve to investigate whether our induction of associations generalizes to new ambiguous and unambiguous words (primes in the EP paradigm).
Hypotheses:
(1) The Prime x Target x Induction interaction is moderated by the SOA

- For short (long) SOA we expect an assimilation (contrast) effect in the EP task:
- In the ambiguity = negative and clearness = positive condition, we expect the following prime × target interaction: if an ambiguous prime precedes a negative target, the reaction time should be faster for short SOA and slower for long SOA than if an unambiguous prime precedes a negative target. If an ambiguous prime precedes a positive target, the reaction time should be slower for short SOA and faster for long SOA compared to an unambiguous prime preceding a positive target.
- In the ambiguity = positive and clearness = negative condition, the prime × target interaction should indicate the reversed pattern for both SOA.

(2) The Prime × Target × Induction x SOA interaction is further moderated by the novelty of the stimuli

- For stimuli used in the induction phase, the four-way-interaction postulated in (1) should be stronger than for new stimuli.

3) Describe the key dependent variable(s) specifying how they will be measured.
Latencies in the EP paradigm constitute the key dependent variable.

4) How many and which conditions will participants be assigned to?
2 (Induced Association: ambiguity=positive and clearness=negative vs. ambiguity=negative and clearness=positive) × 2 (prime type: ambiguous vs. unambiguous) × 2 (target type: positive vs. negative) × 2 (stimulus novelty: same as in induction phase vs. different from induction phase) × 2 (SOA: 200 ms vs. 400 ms) design with repeated measures on the last four factors. Additionally, across participants we counterbalance the specific stimulus set used for learning (see above).

5) Specify exactly which analyses you will conduct to examine the main question/ hypothesis.

We will conduct multi-level analyses with random intercepts for the 24 prime words, 24 target words and participants. In the fixed effects model, we regress the logarithmized reaction times in the EP paradigm on prime type × target type × induction × stimulus novelty × SOA and all subordinate four-way, three-way, two-way interactions and the main effects. The stimulus set (A or B) is included as a covariate. In order to bind error variance, we will include a two-way interaction of the target type with the target type of the previous trial. The target type of the previous trial as a main effect is also included.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

Data of participants who did not finish the online study will not be used. All trials in the EP paradigm with reaction times below 300 and above 3000ms and all false classifications will be excluded from analysis.

7) How many observations will be collected or what will determine sample size?

The link to the online study will be published via the mailing list of the university (students and employees) and will be active for exactly one week (168 hours) before it will be deactivated. Based on previous studies, we expect to collect data from about 400 participants within this period.

8) Anything else you would like to pre-register?

-

Appendix I1

Original and German translation of the Multiple Stimulus Types Ambiguity Tolerance Scale–2 (MSTAT–2; McLain, 2009)

Item No	Original	Translation
1	I don’t tolerate ambiguous situations well.	Ich ertrage mehrdeutige Situationen nicht gut.
2	I would rather avoid solving a problem that must be viewed from several different perspectives.	Ich vermeide es eher ein Problem zu lösen, das von verschiedenen Sichtweisen betrachtet werden muss.
3	I try to avoid situations that are ambiguous.	Ich versuche mehrdeutige Situationen zu meiden.
4	I prefer familiar situations to new ones.	Ich ziehe vertraute Situationen neuen vor.
5	Problems that cannot be considered from just one point of view are a little threatening.	Probleme, die nicht nur von einem Standpunkt aus betrachtet werden können, sind etwas bedrohlich.
6	I avoid situations that are too complicated for me to easily understand.	Situationen, die zu kompliziert sind, um sie leicht zu verstehen, versuche ich zu vermeiden.
7	I am tolerant of ambiguous situations.	Ich bin mehrdeutigen Situationen gegenüber tolerant.
8	I enjoy tackling problems that are complex enough to be ambiguous.	Mir macht es Spaß, Probleme in Angriff zu nehmen, die komplex genug sind, um mehrdeutig zu sein.
9	I try to avoid problems that don’t seem to have only one “best” solution.	Ich versuche es zu vermeiden, mich mit Problemen zu beschäftigen, die scheinbar nicht nur eine 'beste' Lösung haben.
10	I generally prefer novelty over familiarity.	Im Allgemeinen bevorzuge ich Neuheit über Vertrautheit.
11	I dislike ambiguous situations.	Ich mag mehrdeutige Situationen nicht.
12	I find it hard to make a choice when the outcome is uncertain.	Ich finde es schwierig eine Wahl zu treffen, wenn der Ausgang unsicher ist.
13	I prefer a situation in which there is some ambiguity.	Ich bevorzuge Situationen, die ein wenig mehrdeutig sind.

Appendix I2

Original and German translation of the subscale "Discomfort with Ambiguity" from the Need for Closure Scale (Webster & Kruglanski, 1994b)

Item No	Original	Translation
1	I don't like situations that are uncertain.	Ich mag keine Situationen, die unsicher sind.
2	I feel uncomfortable when I don't understand why an event occurred in my life.	Ich fühle mich unwohl, wenn ich nicht verstehe, warum ein Ereignis in meinem Leben aufgetreten ist.
3	When I am confused about an important issue, I feel very upset.	Wenn ich ein wichtiges Thema verwirrend finde, bin ich sehr aufgebracht.
4	In most social conflicts, I can easily see which side is right and which is wrong.	In den meisten sozialen Konflikten kann ich leicht erkennen, welche Seite richtig liegt und welche falsch liegt.
5	I like to know what people are thinking all the time.	Ich mag es zu wissen, was die Leute die ganze Zeit denken.
6	I dislike it when a person's statement could mean many different things.	Ich mag es nicht, wenn die Aussage einer Person viele verschiedene Dinge bedeuten könnten.
7	It's annoying to listen to someone who cannot seem to make up his or her mind.	Es ist lästig, jemandem zuzuhören, der sich anscheinend keine Meinung bilden kann.
8	I feel uncomfortable when someone's meaning or intention is unclear to me.	Ich fühle mich unwohl, wenn die Funktion oder Absicht von jemandem unklar für mich ist.
9	I'd rather know bad news than stay in a state of uncertainty.	Ich würde lieber schlechte Nachrichten wissen wollen, als in einem Zustand der Ungewissheit zu bleiben.

Ich erkläre hiermit, dass ich die zur Promotion eingereichte Arbeit mit dem Titel

Investigating the attitude towards ambiguity: Interindividual differences in automatic activations of evaluations of ambiguity

selbständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt und wörtlich oder inhaltlich übernommene Stellen (alternativ: Zitate) als solche gekennzeichnet habe.

Ich erkläre, dass die Richtlinien zur Sicherung guter wissenschaftlicher Praxis der Universität Tübingen (Beschluss des Senats vom 25.5.2000) beachtet wurden.

Ich versichere an Eides statt, dass diese Angaben wahr sind und dass ich nichts verschwiegen habe. Mir ist bekannt, dass die falsche Abgabe einer Versicherung an Eides statt mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft wird.

Tübingen, den ___________________ ______________________________