The relevance and usefulness of the Implicit Associations Test (IAT) in adolescent development, attachment and depression.

Nicolas Lorenzini

Submitted for the Degree of Doctor of Philosophy

University College London

2015

I, Nicolas Lorenzini confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis focuses on the Implicit Associations Test (IAT), as a measure of implicit cognition aiming to tap distortions of cognitive processing by unconsciously held personal and interpersonal attitudes and beliefs. Classically, IAT had revealed delays in processing of information when dissonant attitudes were activated, even when the participant denied these attitudes. It is a popular measure, but there is little scientific consensus about its value and reliably to detect individual differences across a number of domains. This thesis examined the validity if the IAT in a number of clinical contexts attempting to test its clinical relevance and practical usefulness as a psychometric instrument.

The thesis contains 5 empirical studies: 1) Attachment transmission: The test is administered to a sample of mothers with infants of 1 year of age. IAT was used to measure the mother's implicit attitudes towards attachment relationships and parenting in general and its capacity accurately to predict other measures of attachment, parenting, and psychopathology. In this context, the IAT resulted to be of little value. 2) A newly developed version of the IAT with the potential to measure implicit self-esteem was administered to a large sample of adolescents (14-24 y-o) to test the prediction that implicit self-esteem measured with the IAT is robust to age and gender of the sample. 3, and 5) Depression: Three studies aimed to validate the SE-IAT in the context of depression. The first assessed the ability of the SE-IAT to discriminate between depressed and non-depressed patients. The second study, cross-validated the SE-IAT against several other psychometric instruments in this depressed sample. The third study aimed to assess the value of the SE-IAT to predict and monitor individual gains in the psychotherapy for depression. The SE-IAT is useful at calculating discrepancies between explicit and implicit self-esteem to predict internalizing disorders.

TABLE OF CONTENTS

TABLE OF TABLES

TABLE OF FIGURES

Chapter 1: Introduction

Overview of the Present Thesis

This thesis sets out to assess the performance of the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) in different realms of social cognition. The IAT is a versatile computer-based time-reaction test, which attempts to measure various psychological constructs that are believed to exert an influence over cognition and behaviour in an automatic way, without awareness or conscious control. The following chapters will address its reliability, validity and predictive power through five empirical studies on different samples.

Chapter 1 introduces the concept of implicit or automatic cognition, in particular implicit attitudes, their historical and theoretical origins, their measurement, interpretability, and the issues researchers address regarding its usefulness. It then goes on to introduce the Implicit Association Test and gives an overview of its increasingly widespread use. The following five chapters are distinct studies carried out utilising the measure. Chapter 2 presents an empirical study using two IATs in an infant-mother sample. These measures of implicit social cognition attempt at capturing implicit aspects of attitudes towards parenting and attachment, and their influence on parenting behaviour, parental psychopathology and attachment transmission. Chapter 3 introduces a novel version of the Self-Esteem Implicit Association Test (SE-IAT), developed especially for this doctoral thesis, and applies it on a sample of normally developing adolescents, in order to ascertain its reliability and discriminant validity against demographic characteristics. Chapter 4 makes use of this same version of the SE-IAT on a group of depressed patients and a control group, contributing to knowledge about the relationship of implicit self-attitudes with this prevalent disorder. Chapter 5 replicates this study on a larger sample, comprising only depressed patients, to confirm or disconfirm some of the findings of the previous study. Chapter 6 is a longitudinal study within the context of a randomised controlled trial, where the same SE-IAT is administered 3 times to a sample of depressed patients, some of whom were assigned to a psychodynamic psychotherapy for depression, and others to an inactive control group. It will address the changes of implicit self-esteem and depression, and the relationship between those changes as result of therapeutic intervention. Lastly, Chapter 7 provides a general discussion of the findings, limitations of the studies, issues with the measure and proposals for further research.

The value of this thesis must be seen in the light of its providing a necessary stepping-stone in the construction of an edifice of knowledge around this increasingly popular measure, which is not free of polemic interpretations and issues surrounding its validity.

From Explicit to Implicit: The Evolution of the Scientific Study of Attitudes

When beginning a review of the historical evolution in the study of concept of attitudes, it is first necessary to define the concept. However, selecting one definition of attitude is already a titanic task. This is because the concept of attitude is one of the earliest concepts in scientific psychology and as such, several influential authors have had a say on what an attitude is and what its characteristics are.[1] A simple PsycInfo search of the word "attitude" yields 54 different subject headings and more than 70,000 articles. Besides definitions in common use language, for psychology there is a consensus that attitudes are a multidimensional concept that implies: a) a mental state; b) a value (valence), belief or feeling; and c) a predisposition to behaviour (Altmann, 2008). It seems to be an all-encompassing psychological construct, which has a cognitive, an affective and a behavioural component. It has an object (attitudinal object) upon which a valenced, bipolar (favourable or unfavourable) judgement is passed. This broad way of defining attitudes is the source of great variability in the scientific study of the concept, and concurrently a source of disagreements and controversies (Gawronski, 2007). For the scope of this particular review, I will focus on the controversies surrounding the scientific measurement of attitudes.

As we can surmise from the last paragraph, the study of attitudes implies their measurement, and therefore it has been predominantly a quantitative issue, from very early in the 20th Century (Thurstone, 1931; Thurstone & Chave, 1929). The measures developed by early authors to assess attitudes have relied on the conception that an attitude is accessible by a person via conscious introspection. This conceptualisation of attitudes has the necessary consequence that, if researchers want to know about a person's attitude, the way to proceed is to ask that person to report their attitudes. Measures that openly require an individual to inform about his/her attitudes have been

---

[1] *In psychology, an attitude is an expression of favour or disfavour toward a person, place, thing, or event (the attitude object).* Such is the definition found on Wikipedia, which could be regarded as a consensual definition of the concept. The encyclopaedia's entry on "attitude" has been revised more than 550 times since November 2003. At the time of retrieval, the last modification had been carried out 6 days before.

termed explicit or direct attitude measures. This method proved very fruitful in different fields, and it is still the most popular methodology to assess attitudes: during the year 1989, all published psychological research addressing attitudes utilised at least one direct method to assess them (Greenwald & Banaji, 1995). Early direct measurement has helped us to understand attitudes towards different social groups and institutions (Ferguson, 1935), the influence of the media (Thurstone & Peterson, 1933), and famously, the prevalence of prejudice (Katz, Allport, & Jenness, 1931), among many other phenomena (Sherif, 1935).

The reliance on these direct methodologies, however, rendered conclusions unreliable, especially when attitudes were in relation to socially sensitive topics. This was viewed as problematic by many authors, including Gordon Allport (the most notorious figure in attitudinal research), who in 1935 warned: "the direct frontal attack which many psychological inquiries make, provokes to have a merely conventional answer" (p. 824). Consequently in 1950, Donald Campbell recognised that "in the problem of assessing social attitudes, there is a very real need for instruments which do not destroy the natural form of the attitude in the process of describing it" (Campbell, 1950, p. 15). He reviewed the still scarce, novel and partially validated "indirect measures", which he defined as those measures where the attitude being measured is disguised to the participant.

This review was followed by the development of various creative indirect measures that included observing subjects without them knowing, physiological measures, and deceiving participants regarding the true attitude being measured (Steinberg, 2006). Among these measures we find sentence completion tasks, where a person completes an ambiguous sentence and in doing so confers it a valence (e.g. "It seems to me that segregation…"; Campbell, 1950, p. 17); or unstructured techniques that require the subject to create a story based on ambiguous or neutral graphic stimuli, famously the Thematic Apperception Test (TAT, by Murray, 1943; Prochanski in Campbell, 1950) These measures surely dealt with the attempts at controlling responses by participants based on self-presentation biases and social desirability (De Houwer, 2006). Nevertheless by 1989, the same survey mentioned above found that only 6% of published studies on attitudes used indirect techniques to measure attitudes. The authors of this survey argued that the reason for this disparity was the assumption that attitudes operate exclusively in a conscious mode (Greenwald & Banaji, 1995).

It sounds counterintuitive that researchers in a field that had developed highly creative indirect measures of attitudes still believed that they only operated at a conscious level. However, this idea was based on evidence.

Until the 1960s, there was an assumption that a one-to-one correspondence between attitudes and behaviour existed (Fazio, 1986). This conception was rooted in many definitions of attitudes, to the extent that when they this correspondence was not found, it would be logical to think that these were not "true" attitudes. However, correlations between measurements of attitudes and actual behaviour are generally weak (LaPiere, 1934; Wicker, 1969), to the extent that this last author proposed abandoning the research on attitudes to find better and more consistent predictors of behaviour. However, the way attitudinal research progressed was by trying to identify moderators to the relationship between attitudes and behaviours. Investigations in this field focused on the contextual factors moderating this relationship, and the mechanisms of this moderation, or the *When* and *How* attitudes would predict behaviour (Fazio, 1986; Fazio, Powell, & Williams, 1989). The results of these efforts could be summarised by the finding that attitudes would predict behaviour when the former are strongly activated and/or when the person clearly perceives a link between attitude and behaviour (Greenwald & Banaji, 1995).

These last authors convincingly argue that the process of finding the specific circumstances and conditions in which an attitude would predict behaviour had logically reduced the scope of the concept of attitudes. They take multiple examples from scientific experiences accrued until that time, which seemed to justify the existence of attitudes of which the person was not conscious, yet they are strongly predictive of behaviour. This, coupled with the interest stirred up by findings in implicit memory, led them to propose a novel way to conceptualise attitudes, and in the process, social cognition as a whole.

Implicit Social Cognition: Its Origins and Status within Psychological Science

Implicit social cognition is a term coined by Greenwald and Banaji (1995), to include all those traces of past experience affecting some social cognitive performance, when the earlier experience is unavailable to self-report or introspection. In their seminal paper, they warn that they are not creating anything new, but subsuming a series of hitherto loosely connected, but readily observable phenomena in social cognitive psychology. They took advantage of the increasing interest in implicit

memory at the time, and even borrowed the term "implicit" from the group of investigations relating phenomena in social cognition whose definitions did not include an unawareness component, but whose experimental results warranted one. This had a profound consequence on the way that cognitive phenomena were interpreted and more importantly, brought about a fruitful period in the development of measures attempting to capture the "implicit" aspects of social cognition, which up to that point were operationalised in terms of responses given by subjects in direct, self-report measures.

Attitudes, being the most "distinctive and indispensable concept" of social psychology (Allport, 1935, p. 798), were the first in line for this change of paradigm. As Greenwald and Banaji noticed (1995), the definitions of attitudes at the time did not preclude an implicit component. They cite influential definitions of attitude spanning a period from 1931 (like Thurstone's: "attitude is the affect for or against an object", cited in Greenwald & Banaji, 1995) to 1962 (Krach, Crutchfield and Ballanchey's: "enduring systems of positive or negative evaluations, emotional feelings, and pro or con action tendencies with respect to social objects"; p. 7). Those theoretical caveats, added to the widespread idea that attitudes would only predict behaviour when a strong conscious component was present, left the scientific field of social cognition in a state of contradiction. Social cognitive scientists kept finding automatic effects in social cognition which acted on a subliminal level (like the classic "halo effect"; Thorndike, 1920), and many influential researchers argued against the reliance of the psychological sciences on verbal self-report, given the obvious poverty of subjective introspection (Nisbett & Wilson, 1977). It seemed that cognitive psychologists were completing a historical full circle, returning to conceptions of a mind working outside consciousness, which could be as effective on behaviour as the conscious mind. Were they, in laboratory settings, getting closer to what psychoanalysts had discovered in the consulting room?

The Influence of Psychoanalysis and its Epistemological Issues

Certain schools of thought within psychology have been founded on the hypothesis that much of mental life occurs outside subjective awareness (Freud, 1916). However, the lack of a systematic way of testing these hypotheses beyond the experiential clinical realm caused a loss of popularity in these theories in the ambit of academic psychology, which is slowly being overcome. Unconscious mental processes are increasingly being studied by academic cognitive and social psychology approaches, especially regarding the concept of "attitude", which has been traditionally defined as

having a cognitive component and an affective-evaluative component (Allport, 1935). Attitudes that are outside awareness could be automatically activated without cognizance or volitional control, a phenomenon known to cognitive psychologists as "implicit attitudes". This model of attitudes seems to converge with the psychoanalytic construct of an unconscious mind, which continuously acts upon mental life as a whole (Luborsky & Barrett, 2006; Westen, 1998). Yet, psychoanalysis and cognitive theories of the mind have developed in parallel to each other as alternatives, making it difficult to directly compare an empirical/quantitative conception of an unconscious mind, with the psychoanalytic conception, based on clinical experience and hermeneutics.

The evolution of these parallel conceptions started with behaviourism, which developed in psychological academia as an alternative to psychoanalysis, and dominated this field during the 1950s, especially in the United States. It rejected the notion that mental processes (conscious or unconscious) could play a causal role in behaviour. During the 1960s, its protagonism within academia was taken over by cognitive experimental psychology. Cognitive psychology saw the mind as a serial processor of information, where memories of past events only become significant when they enter short-term memory (or working memory, which can be loosely understood as consciousness). This conception guided most of the work of cognitive psychologists until the late 1980s (Westen, 1998). However, as we have reviewed above, such a model of the mind eventually proved inadequate to explain most functions of the mind and to predict behaviour (Greenwald & Banaji, 1995).

In the late 1980s, Kihlstrom (1987) published a very influential paper inaugurating the existence of the "cognitive unconscious" (p. 237). This author highlights the obsolescence of a classical (serial) information processing model of the mind, which implies the engagement of attention (working memory, consciousness) as a pre-requisite for a cognitive analysis of mental content. This type of model cannot readily explain observable phenomena such as the acquisition and use of procedural knowledge, hypnosis, priming, or subliminal perception. Novel cognitive models of the mind, such as the parallel distributed processing or connectionism, assume that when a concept is activated within the mind, associated nodes can also become activated automatically outside awareness and influence experience, thought, and action. Such a concept had been advanced by classical thinkers in psychology, such as Freud (1915), Helmholtz, James and Janet, who in turn had taken it from the philosophical zeitgeist of the German idealism of the 19th century (Kihlstrom, 1987; Kihlstrom, Barnhardt, & Tataryn, 1992; Kihlstrom, Mulvaney, Tobias, & Tobis, 2000).

The conception of the unconscious at which experimental psychology arrived was, in the beginning, a purely cognitive unconscious. It did not include affective or motivational elements, but it was soon discovered that affects and motivation could also function outside awareness in experimental settings. This more complete unconscious is known as the "psychological unconscious" among experimental psychologists (Kihlstrom et al., 1992). This psychological unconscious, which includes not only thinking operations, but also affective and motivational elements, resembles the psychoanalytic version of the unconscious. Nevertheless, it is extremely difficult to equate both conceptions of the unconscious mind, given their diverging intellectual origins. In contrast to psychoanalysis, the unconscious of cognitive experimental psychology is chiefly a cold, automatic, and almost exclusively cognitive feature of the mind. The fact that it is outside awareness is only by virtue of mental architecture, and not because of its conflictive nature (Epstein, 1994; Masling, 2000; Westen, 1998). On the other hand, the psychoanalytic unconscious is linked to primitive, instinctual and emotional processes. This implies that it fails to distinguish other kinds of unconscious processes (especially cognitive processes), and to recognise that many unconscious processes can be adaptive and learned (Westen, 1998). However, the new psychological unconscious of cognitive approaches bears a strong resemblance to the psychoanalytic unconscious, especially that of contemporary psychoanalytic approaches in search of integration (Fonagy, 1982; Luyten, Blatt, & Corveleyn, 2005; Modell, 2008). Contemporary psychoanalytic theorists view the unconscious as a cognitive-affective-motivational feature, increasingly disengaging from drive theory and sexuality, particularly in North America and Britain (Budd, 2001). Nevertheless, the equation of these two concepts of the unconscious is far from straight-forward: the unconscious described by academic psychology which emerged a century later comes from a different theoretical tradition and a different set of methods.

The problems encountered when pairing concepts from cognitive psychology and psychoanalysis are not circumscribed to the unconscious only. Traditionally, while psychoanalysis has always been scientific in its aim, its methods have grown apart from positivist science (Kandel, 1999; Masling, 2000). In some psychoanalytic circles, the mere suggestion of empirically testing the theory's tenets meets much resistance, as if these tenets were dogmas to be protected more than assertions about nature that could be falsified (Blass, 2010; Blass & Carmeli, 2007; Kihlstrom, 1994; Westen, 1998). However, it would not be correct to hold these circles responsible for the distancing between psychoanalysis and academic psychology. As with all social sciences, the

processes used to collect and produce data shapes the latter somewhat (Diesing, 1985). Traditionally, psychoanalysts have relied on evidence gained from clinical interactions in the consulting room. These interactions are the basis for making inferences regarding the validity of psychoanalytic notions of underlying personality theory, psychodynamics, treatment strategies, and for generalising from the individual case study to the general population (Masling & Cohen, 1987).

Such a method of collecting data makes psychoanalysis a relatively configurational science (in contrast to a correlational science). Psychoanalytic theories are not composed of statements correlating pairs of variables (Kihlstrom, 1994). Variables and hypotheses are present, but they are not tested with a causal or correlational framework (Diesing, 1985). That is to say, psychoanalysis has a very complex object of study, which exhibits more variability and a wider range of properties to be considered than empirical cognitive psychology. The latter is a more restricted science, where we can find a strong tendency to view only those properties that can be located as discrete points on an arithmetic continuum. From a configurational standpoint, qualitative differences become much more important (Whitley, 1978). For example, academic psychology did not develop a concept such as over-determination (Laplanche & Pontalis, 1988, p. 292), central to psychoanalytic theory, probably because its methodologies tend to exclude rather than to reveal multiple determination of psychic phenomena. However, the problem is still implicitly present in their conceptualisations: every behaviour has perceptual, learning, cognitive, motor and other components, and mathematical models of multiple causation (such as multiple regression models) have been largely used by academic psychology to explain outcomes based on the discrete contribution of multiple causes (Kruglanski & Stroebe, 2012).

The same can be said about the notion of unconscious processes. While psychoanalysis strives to find them in every corner of mental life, cognitive experimental psychology has had to recognise their existence when explanations about the hegemony of consciousness were not adequate enough to cover the observable, as we will see in the next section.

The Emergence of Implicit Attitudes: Empirical Results that Pointed Toward "Something Other" than Consciousness

As briefly mentioned above, implicit social cognition took advantage of the robustness of findings on implicit memory (also known as priming) and the interest these elicited at the beginning of the 1980s. Implicit memory is basically the evidence

that the exposure to information that cannot be consciously remembered facilitates the performance of related information that is presented later (Tulving, Schacter, & Stark, 1982). This happens even when the previous information has been presented subliminally, or to patients with amnesia (Tulving & Schacter, 1990). These authors revolutionised the field of memory when they discovered that there was something other than declarative and semantic memory: "… since there are problems with [this results'] interpretations in terms of modifications of semantic memory, we are tempted to think that they reflect the operation of some other, as yet little understood, memory system" (Tulving et al., 1982, p. 341). These memory systems were rapidly used to study social cognition (e.g. gender stereotyping, Banaji, Hardin, & Rothman, 1993)[2].

Also, as mentioned above, halo effects were noticed in experimental settings as early as 1920 (Thorndike). This effect involves the transference of the valence of a positive attribute of an object to an unknown attribute, without a logical association between the two attributes. A classic example is a study in which students remembered an academic's height depending on how the academic was introduced to them: as a lecturer, senior lecturer or professor. Students remembered the person to be taller depending on their status (Wilson, 1968). If directly asked, students were unable to recognise that bias. But as with most implicit social cognitive effects, once the manipulation is explained, the effect dissipates (Greenwald & Banaji, 1995).

This same phenomenon occurs only by mere exposure to a stimulus. People will prefer a neutral stimulus from that to which they have been exposed before, even subliminally. Once again, if participants are informed of this effect, the effect disappears (Bornstein & D'Agostino, 1992; Zajonc, 1968).

All these phenomena suggest that there must be some sort of unaware or unconscious cognitive processing of valences towards an object, and that this processing is predictive of behaviour. Therefore, it was valid for Greenwald and Banaji to coin the term "implicit attitudes" (Greenwald & Banaji, 1995, p. 9) to cover this whole span of phenomena. These effects can also be found in attitudes towards the self, which will be more thoroughly reviewed in subsequent chapters, such as preferences for one's own group, even if the group one pertains to is artificial or experimentally created

---

[2] Priming methods are still used, yet highly criticised for their lack of replicability. It is not the scope of this review to dwell on this criticism, but it might be enough to say that the scientific panorama is still divided when it comes to the usefulness of priming models and their explanatory theories (Cesario, 2014b; Janiszewski & Wyer Jr, 2014).

(Brewer, 1979; Tajfel, 1974); or the fact that one is more likely to be persuaded by a side of an argument if one has to defend it, even if the allocation of sides has been openly random, and even if the argument is not necessarily in accord with private convictions (Janis & King, 1954); or the fact that one tends to prefer a stranger of whom it is said that they hold opinions similar to one's own (even in irrelevant topics), and they are judged as more intelligent than another stranger who does not hold similar opinions (Byrne, 1961).

Following the establishment of implicit social cognition in the mid-1990s this scientific field was, for a second time, benefited by a plethora of novel and creative instruments to measure individual differences in implicit attitudes, in order to reliably ascertain their origins and causes, and their relationship to other psychological constructs and to behaviour (Koole et al., 2007).

Measuring Implicit Attitudes

Examples of instruments measuring implicit attitudes are usually based on existing techniques measuring other cognitive traits. For instance, the evaluative priming technique has been used to measure racial attitudes (Fazio, Jackson, Dunton, & Williams, 1995). Participants are presented with photographs of black or white faces (the prime) and then with positive or negative words (the target). Participants are asked to rapidly recognise the valence of the word presented. Facilitation effects can be seen (in white participants) for the trials in which a black face is presented before a negative word and white faces are presented before a positive word. These facilitation effects are taken as a measure of implicit racial attitude towards black and white people.

A similar procedure presented primes subliminally, to measure implicit gender attitudes. Primes were positive or negative words and targets were male or female names in one task. In another task male or female names were the primes and the targets were positive or negative words. Again, there were noticeable priming effects that were interpreted as attitudes towards gender (Draine & Greenwald, 1998).

In order to measure the automaticity of emotional valences given to an object, De Houwer and Eelen (1998) developed an affective variant of the Simon paradigm. The original Simon task (Simon, 1990; Simon & Rudell, 1967) required participants to press a left key on a keyboard when a red light was presented, and a right key when a green light was presented. They discovered that the time that elapsed from the presentation of the light to the participant's response was increased when the light (regardless of its colour) was presented on the opposite side of the response key, i.e. responses to the red

light will be slower if the light was presented on the right side. This paradigm, purely perceptual and procedural, relied on a task-relevant trait of the stimulus (the colour of the light) and an irrelevant trait (the position of the light on the visual field of the participant), to find that the task-irrelevant trait affects task performance. The affective variant utilises word-stimuli that are positive (flower), negative (cancer) or neutral (paper). These words could be either nouns (as above) or adjectives (honest, stupid or normal, respectively). The instruction to participants is to say "positive" when the word is a noun, and "negative" when it is an adjective. The actual valence of the word presented is task-irrelevant and should be ignored. However, they found that performance is impaired when the actual (task-irrelevant) valence of the word is incompatible with the response. This type of task lends support to the automaticity of emotional valence processing (De Houwer & Eelen, 1998).

Another example is the Go/No-Go Association Task (GNAT; Nosek & Banaji, 2001), in which participants must rapidly press a key when they see a stimulus appear that belongs to one of two categories permanently present on the screen. Stimuli can be fruits, positive words, negative words, and insects. Participants are given a very short response window. The difference in accuracy (in other words, how many times the participants correctly and rapidly press the key when a stimulus is presented) versus their mistakes or lack of response in counterbalanced tasks (i.e. when they must press the key when fruit and positive must be classified vs. when fruit and negative words must be classified) is taken as a measure of the automatic attitude towards fruit. In short, if a person makes fewer mistakes when classifying both fruits and positive words than when classifying fruits and negative words, a positive attitude towards fruits is assumed. This task has been used to measure racial and gender attitudes, showing a general positive attitude towards females and towards white race. These results were only partially correlated to explicit measures of racial and gender attitude.

These and other measures are increasingly popular in psychological and clinical research (Steinberg, 2006). They have been used to measure racial attitudes, self-esteem (Hetts, Sakuma, & Pelham, 1999), implicit attitudes towards product brands (Wanke, Plessner, Gartner, & Friese, 2002), presidential candidates (Nosek, Banaji, & Greenwald, 2002), and stereotypes (Fazio et al., 1995).

The Present Thesis: The Implicit Association Test (IAT)

The most popular of these measures is the Implicit Association Test (Greenwald et al., 1998). It is a very versatile time-reaction computer-based test, which is designed to measure the differential strength of the association between a target-concept dimension and an attribute dimension in comparison with a contrasting concept and an opposite valenced attribute. Here I provide a brief and general description of the test, with more detailed descriptions being given in subsequent chapters to highlight the different versions of the IAT used in the studies for this thesis.

The first experiment published using the IAT (Greenwald et al., 1998), measured the strength of the associations between flowers and the pleasant attribute, and insects and the unpleasant attribute, which are associations that can be expected in the general population (with the improbable exception of participants being both entomologists or allergic to flowers). The IAT comprises 5 discrimination tasks (Figure 1). The procedure begins with the introduction of the target concept. In the case of Figure 1, the first task is to press a left key for names of insects and a right key for flower names. The second step is the introduction of the attribute dimension, categorising words as pleasant or unpleasant in meaning. After these two introductory steps, both the target-concept and the attribute are superimposed in the third step, in which stimuli for concept and attribute are presented to participants in alternated trials. In the fourth step, the participants learn a reversal of the key assignments for the target-concept, and the fifth and final step combines the attribute discrimination (not reversed) with the reversed target discrimination.

It is expected that the participant will find one of the combined tasks more difficult than the other combined task (steps 3 and 5). The measure of this difficulty is the difference in time-reactions (latencies) between steps 3 and 5. In the original IAT, it was expected that participants would take longer to respond to trials in step 5 than trials in step 3. Such latency difference provides a measure of implicit attitudinal difference between target categories (between flowers and insects). The algorithm to calculate this measure of relative attitudinal strength and other procedural details are explained in Chapter 3.

Uses

The IAT is very versatile. This same configuration of trials and steps can be used to measure implicit attitudes towards practically any attitudinal object, with regards to a contrasting object. There is a large amount of research both relating implicit attitudes to behaviour and psychopathology, improving the procedure, and scoring and interpreting the measure (Aberson & Beeney, 2007; Greenwald & Nosek, 2001; Greenwald, Nosek, & Banaji, 2003; Greenwald, Poehlman, Uhlmann, & Banaji, 2009; Lane, Banaji, Nosek, & Greenwald, 2007; Nosek, Greenwald, & Banaji, 2005, 2007a; Schnabel, Asendorpf, & Greenwald, 2008).

Among the various adaptations of the IAT in studies using this measure we can

Figure 1: Schematic description and illustration of the Implicit Association Test (IAT). The typical IAT procedure, used throughout this thesis, involves a series of five discrimination tasks (numbered columns). A pair of target concepts and an attribute dimension are introduced in the first two steps. Categories for each of these discriminations are assigned to a left or right response, indicated by the black dots in the third row. These are combined in the third step and then recombined in the fifth step, after reversing response assignments (in the fourth step) for the target-concept discrimination. The illustration uses stimuli for the specific tasks for one of the task-order conditions of the first published IAT (Greenwald et al., 1998), with correct responses indicated with black dots, in the fourth row.

Sequence	1	2	3	4	5
Task description	**Initial target-concept discrimination**	**Associated attribute discrimination (compatible task)**	**Initial combined task**	**Reversed target-concept discrimination (incompatible task)**	**Reversed combined task**
Task instructions	• INSECT FLOWER •	•pleasant unpleasant•	•INSECT •unpleasant FLOWER• pleasant•	INSECT• •FLOWER	INSECT• pleasant• •FLOWER •unpleasant
Sample stimuli	poppy • •ant rose• •bee orchid• •tarantula	•freedom abuse• •health grief• •love poison•	freedom• •bee orchid• love• •grief •poison rose•	•poppy ant• •rose bee• •orchid tarantula•	freedom• bee• •orchid love• •grief •poison •rose

find some as dissimilar as measuring implicit attitudes towards meat and vegetables in meat eaters and vegetarians (Barnes-Holmes, Murtagh, Barnes-Holmes, & Stewart, 2011), the spontaneous generation of affect in everyday life (Conner & Barrett, 2005), predicting dental flossing behaviour (Millar, 2011), measuring distress caused by tinnitus (Moring, Bowen, & Thomas, 2014), and implicit self-esteem in humour styles (Stieger, Formann, & Burger, 2011) and in schizophrenic patients (Valiente et al., 2011). As we will see in the next chapter, the IAT has also been used to predict adult attachment styles (Dewitte, De Houwer, & Buysse, 2008), and other personality variables, such as neuroticism in relationship with the Big Five Model. It has been shown that implicit neuroticism is an incremental predictor to self-report, of more than 50 behaviours (Back, Schmukle, & Egloff, 2009). There is a child-friendly version of the measure, which revealed a negative stereotype towards black faces very early (6 years old), comparable to the implicit attitude in adults. Interestingly, what differentiates children's attitudinal tendencies from those of adults' is that the latter show a decrease in explicit negative attitudes towards black people (Baron & Banaji, 2006). Children tend to show an increased in-group bias unless externally controlled by an adult (Rutland, Cameron, Milne, & McGeorge, 2005).

Being a measure created within the framework of social cognitive psychology, much of the use of the IAT is to be found in topics such as social stereotyping and prejudice towards various groups. The creators of the measure maintain a website[3] where people can complete IATs online on a broad spectrum of attitudinal topics: gender stereotypes, race stereotypes (Arab-Muslim, Jewish, Black, Oriental, Native American), attitudes towards older people, attitudes towards US presidential candidates, towards disabled people, towards homosexual people, and towards overweight people (for an extensive review, see Nosek et al., 2002; and Nosek et al., 2007b).

With regards to psychopathology, the use of the IAT has yielded interesting results. For example, an IAT measuring the association between the self and alcohol consumption, i.e. a measure of alcohol identity, is able to predict engagement in risky college drinking practices after a 3 and 6 months follow-up, even when controlling for standard alcohol consumption measures (Gray, LaPlante, Bannon, Ambady, & Shaffer, 2011). The IAT has been found to predict binge drinking in adolescents (Thush & Wiers, 2007; Thush et al., 2008). Along the same lines, an IAT measuring implicit

---

[3] https://implicit.harvard.edu/implicit/

attitudes towards drugs was capable of predicting relapse in heroin addicts undergoing detoxification and addiction treatment (Marhe, Waters, van de Wetering, & Franken, 2013). In eating disorders, it has been found that implicit self-esteem as measured with the IAT is higher in bulimic patients than in healthy controls, which is in contrast to a comparatively lower explicit self-esteem (Cockerham, Stopa, Bell, & Gregg, 2009). Such a pattern has also been found in young women with depression, suicidal ideation and loneliness (Creemers, Scholte, Engels, Prinstein, & Wiers, 2013). Bipolar patients show increased depressive self-associations compared to controls, at a similar level as unipolar depressed patients (Jabben et al., 2014). Manic patients show higher implicit self-esteem, as measured with the IAT than euthymic patients and a trend towards higher self-esteem than healthy controls (Park et al., 2014). The use of the IAT on mood disorders, in specific depression, will be more deeply reviewed in Chapter 4.

In measuring implicit self-esteem, it has been found that patients with social anxiety have a significantly lower implicit self-esteem than healthy controls, and that patients with post-traumatic stress disorder (PTSD) have a distinct level of implicit self-esteem between those two groups (Glashouwer, Vroling, de Jong, Lange, & de Keijser, 2013). An independent group replicated that result in the same year (Ritter, Ertel, Beil, Steffens, & Stangier, 2013). In the case of PTSD, patients who currently have the disorder show lower implicit self-esteem than remitted patients (Roth, Steffens, Morina, & Stangier, 2012). Patients with schizophrenia also show decreased levels of implicit self-esteem as measured with the IAT (Valiente et al., 2011), but among this group, implicit self-esteem levels are similar for delusional and non-delusional patients (Nakamura et al., 2015). In the particular case of the narcissistic personality disorder, the IAT measuring implicit self-esteem has been useful to cast doubts on the widespread theory of fragile self-esteem in narcissism, which postulates that a grandiose self actually hides and protects implicit low self-esteem and a diminished sense of self (Emmons, 1987; Morrison in Lima, 2007). While Lima (2007) was incapable of finding low implicit self-esteem in narcissistic patients using the IAT, further investigations with the same measure found that higher implicit self-esteem was positively associated to severity of narcissistic manifestations (Vater et al., 2013).

The IAT has also been used to measure self-stigma in mental health, i.e. the extent to which people affected by mental health problems internalise negative stereotypes about themselves. Operationally, implicit self-stigma is the product between measures of negative attitudes towards mental health and low implicit self-esteem. Such a variable was able to predict the quality of life of participants independently of

diagnosis, depressive features and demographic variables (Rüsch, Corrigan, Todd, & Bodenhausen, 2010).

Most of the clinical investigations reviewed in the last paragraph have a cross-sectional design. However, it is important to note that in spite of having good indexes of test-retest reliability, the IAT is also malleable, and to an extent, context dependent (Blair, 2002; Blair, Ma, & Lenton, 2001). For example, implicit preferences for white or black faces can be affected if before the administration of the IAT, participants are exposed to photos of admired black persons (Dasgupta & Greenwald, 2001). Moreover, IAT effects are different in depressed patients if, before completing a version of the Self-Esteem IAT, they are subjected to mood induction techniques (Gemar, Segal, Sagrati, & Kennedy, 2001). Such malleability is useful, for example, to measure changes in implicit attitudes as an outcome of psychotherapy (De Houwer, 2002). For example, a study reports changes in implicit associations between social situations and anxiety, as measured with the IAT, after conditioning social situations to positive stimuli. Patients who went through this conditioning intervention were more likely to complete an impromptu speech in front of the experimenters (who also told the participant that other people would rate their performance recorded on video), although they did not report less explicit anxiety during the completion of the task (Clerkin & Teachman, 2010). A similar study showed reductions in both implicit and explicit anxiety after group treatment of social phobia (Gamer, Schmukle, Luka-Krausgrill, & Egloff, 2008). In spite of these results, given that the malleability of implicit attitudes seems to be only partial, and comparable to measures of explicit self-esteem (Buhrmester, Blanton, & Swann, 2011), many studies show that remitted patients show positive changes in the implicit attitudes measured, but not to the level of controls without a history of psychopathology (Glashouwer & de Jong, 2009).

In sum, the IAT has helped to produce large quantities of research in all corners of psychology, given its versatility and the little amount of effort and time required from participants. It can be added to mostly any research design, without major considerations about the burden placed on participants. However, in spite of (or maybe because of) its widespread use, interpretation of its results is still a matter of discussion, and there is no single theoretical framework with which to understand them.

Theoretical Framework Explaining IAT Effects

What is clear about the IAT, and implicit cognition in general, is that it has come to provide the *coup de grâce* to cognitive models of serial processing. There are at least two modes of information processing in the mind, and these are interrelated but to some extent independent. However, the way that they relate, in order to give rise to attitudes, and consequently behaviour, has been addressed by various theoretical models specific to attitude functioning. These models belong to a group of theoretical explanations called "dual models" of cognition (Fazio & Olson, 2003, p. 301)

One of these models is Epstein's Cognitive-Experiential Theory (CEST; Epstein, 1994). This theory posits that individuals' responses to the environment are organised on two distinct levels: one rational and one experiential. The rational system requires deliberate effort and motivation. Its functioning is declarative, conscious and therefore it is slow. The second system, which is experiential, processes information rapidly, without conscious control. These two systems interact, so behaviour is influenced by both. In the case of attitudes, explicit evaluations rely mainly on the rational processing, while implicit attitudes are chiefly under control of the experiential system (Epstein & Morling, 1995). Epstein compares the experiential processing to Freud's pleasure principle (Freud, 1916), but criticises its lack of an evolutionary justification, because according to Epstein, for Freud this principle only gives way to symptoms, dreams and slips (Epstein, 1994). This critique is only partially justified, because while the pleasure principle is indeed viewed as a primitive troublemaker during adult social life, its evolutionary advantage during the first months of life sets the basis for subsequent mental development and attachment (Bowlby, 1969; Freud, 1914). The functioning of the experiential system is "associationistic" (Epstein, 1994, p. 718), where concepts are linked through affect (including wishful thinking, concrete visual imagery or metaphors, displacement, condensation and lack of considerations of the constrains of time and space), while in the rational system, concepts or schemata are linked through linguistic associations and conventional rules of logic.

More specifically regarding attitudes, the Dual Attitudes Model (Wilson, Lindsey, & Schooler, 2000) argues that implicit attitudes are those that have been replaced by new attitudes. It is not that a new attitude replaces an older one, but only overrides it. Among these two attitudes, the one that will predict behaviour or endorsement by the subject, will be defined by the cognitive capacity to retrieve the explicit attitude. Older (implicit) attitudes have been automatised, and they are prone to have been overlearned.

Thus they emerge spontaneously when an individual is faced with an object, and only cognitive capacity and motivation can cause the emergence of explicit attitudes. This way of distinguishing implicit "automatic" attitudes from explicit "reflective" ones is problematic, because newly learnt attitudes can also been activated automatically (De Houwer, 2006; De Houwer Dirk Hermans Paul Eelen, 1998).

One of the most popular theories explaining implicit and explicit processes stems from an earlier study of the relationship between attitudes and behaviour (Fazio, 1990). The MODE model (Motivation and Opportunity as Determinants) posits that the magnitude of the relation between implicit and explicit attitudes and their influence on behaviour is contingent to the individual's motivation and opportunity (time and resources) to deliberate (Fazio & Olson, 2003). This theory implies that cognitive processes are never purely spontaneous nor deliberative (De Houwer, 2006; De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009).

Following this same line, and more akin to the procedural characteristic of the IAT, the Reflective-Impulsive Model of social behaviour (Strack & Deutsch, 2004) proposes that perception, thinking and behaviour are functions of two systems: the Reflective and the Impulsive systems. The reflective system is propositional, in the sense that the reasoning process in this system, for example about the self, is delivered as a propositional decision, fully mediated by language, as in *"I am happy"*. This reasoning process includes concepts (*I* and *happy*), linked by a linguistic relation (*am*). As with any proposition, it can be regarded as either true or false, and is generated by introspection. The impulsive system processes information via an associationist spread of activation. That is to say, concepts are not connected by a falsifiable link, but by an affective association. An impulsive attitude does not need the link "am", because there is an automatic activation of the concept "happy" when the concept "I" is activated (Schnabel et al., 2008). In this sense, implicit attitudes gain in strength when two concepts share the same affective valence.

What these theories have in common, is that they show that implicit and explicit attitudes are rarely found in isolation. Behaviour can then be explained by an interplay of these levels. These ideas are supported by neuroscientific evidence. It is known that the amygdala has a central role in the recognition of emotion, being one of the first areas in the brain to be activated in an emotional state. Note that emotional activation is, most of the time, outside conscious control, and implicitly produces autonomic and endocrine responses, even if the object eliciting the affective response is presented

subliminally (Cunningham et al., 2004; Phelps & LeDoux, 2005). During the administration of a race IAT, (white) participants in an fMRI scanner showed increased activation of the amygdala when presented with black faces. This activation was correlated with the IAT results, but not with explicit self-reports on racial attitudes. The activation of the amygdala diminished when the black faces shown belonged to familiar and positively regarded black individuals (Phelps et al., 2000). Familiarity with novel faces, achieved by repetition of stimuli, provoked a decrease in amygdala activation, but that decrease was significantly faster for white than for black faces (Hart et al., 2000). Here we see a cerebral activation that is known for its automatism, speed and relative unavailability for introspection, which is captured by the IAT. It is noteworthy that amygdala activation is stronger when the stimulus is presented subliminally, thus with reduced conscious control, and reduced top-down inhibition of the amygdala by cortical structures (Cunningham et al., 2004). Given the time-reaction paradigm of the IAT, it seems that the longer time taken by respondents during incompatible trials reflects the time taken by cortical structures to inhibit amygdalar activation. Indeed, there is evidence that the dorsolateral prefrontal cortex and the anterior cingulate cortex regulate amygdala response during racial IATs in fMRI designs (Chee, Sriram, Soon, & Lee, 2000; Cunningham et al., 2004; Stanley, Phelps, & Banaji, 2008). A study within an EEG design showed that these inhibitory responses do indeed take a longer time, between 600 and 700 ms (Egenolf et al., 2013).

In sum, from both a theoretical standpoint and from neuroimaging evidence, it is clear that behaviour and attitudes are constructs that require analysis on at least two different levels of information processing. Given that implicit attitudes are connected to automatic valuations and spontaneous responses, with limited volitional control and awareness, measuring these attitudes appears to be important for those behavioural phenomena that appear outside the participant's control and, in specific, for clinical phenomena.

Criticism, Inconsistencies and the Need for Further Research

Despite the large amount of research in a vast spectrum of psychological enquiry, the IAT has not been exempt from criticism. There are critics of the whole spectrum of implicit measures, critics of the IAT as a group of measures and specific critics of particular versions of the IAT. I will refer to the first two critiques in this section. The

inconsistencies and controversies regarding specific versions of the IAT will be reviewed in the chapters in which they are addressed in more detail.

Regarding criticism towards implicit measures in general, many authors refer to construct validity. In particular, there is a questioning of the capacity of these measures to actually assess "unconscious" features. Without necessarily referring to the dynamic unconscious of psychoanalytic theory, many critics of implicit measures doubt their capacity to measure attitudes that are outside awareness. Definitions of implicit measures usually include this unawareness component, however they do not specify what kind of awareness they are referring to. Attitudes have at least three loosely connected aspects by which they can be considered outside awareness (Gawronski, Hofmann, & Wilbur, 2006). There is source awareness - meaning that the individual often lacks awareness of the cause of their attitudes - as seen in the mere exposure effect (Bornstein & D'Agostino, 1992; Zajonc, 1968); content awareness, which is the awareness of the attitude itself and the possession of such attitude, as is observable in many prejudice studies in which participants sincerely cannot recollect prejudiced attitudes towards social minorities; and impact awareness, which is knowledge about the impact that an attitude has over other psychological processes.

Regarding source awareness, Olson and Fazio (2001) developed a classic conditioning paradigm that paired a hundred randomly chosen words with either negative or positive unconditioned stimuli. After conditioning, an IAT yielded stronger positive associations for those words that had been paired with the positive unconditioned stimulus. Upon enquiry, participants declared they were unaware of the pairing of these words with a positive stimulus. This study was later replicated with a priming task instead of an IAT (Olson & Fazio, 2002). However, source awareness is not exclusive of implicit measures. In the case of the mere exposure effect, individuals can declare they prefer an alternative (to which they had been exposed) over another, but cannot explain the reasons for such a preference.

There is consistent evidence that much of what is measured by implicit instruments is not devoid of content awareness. The literature tends to show that implicit measures and their explicit counterparts are positively correlated, even if weakly (Greenwald et al., 2009; Lane et al., 2007; Nosek et al., 2007a). This implies that the content of the implicit attitude is already present in consciousness. In support of this idea, various studies have shown that when participants are asked to portray themselves honestly in self-report measures of self-esteem, or when they are told that

experts will be able to identify if someone is lying in a self-report on race, these correlations become stronger (Nier, 2005; Olson, Fazio, & Hermann, 2007).

The lack of impact awareness is observable when the activation of an attitude has effects on other psychological processes, even when there is motivation and enough cognitive resources to control for such effect. Gawronski et al (2006) found that German participants rated neutral behaviours more negatively when these were performed by Turkish people than when performed by German people. The IAT moderated this evaluative tendency, but not their explicit attitudes nor their motivation to control for prejudiced reactions (which in turn moderated only self-report). In this case, participants were aware of the racial content of the evaluating task (they knew that assigning valence to behaviour could be reflecting their racial tendencies towards Turkish people, and they had all the time they wanted to assign a valence to these behaviours), but in spite of this explicit knowledge, implicit attitudes still influenced their responses. These authors conclude that the only "unconscious" feature of implicit measures is the lack of impact awareness.

These types of unawareness refer to the attitude being measured, rather than to the process of measuring it. Regarding the capacity of respondents of the IAT to be aware of what is being measured, many investigations coincide in showing that people are sometimes aware of what is being measured, just immediately after the test or even during its administration. In a black-white racial IAT, 64% of respondents realised they were slower during the incompatible task. Of these, 37% attributed this difficulty to the fact that they apparently had more negative attitudes towards black people (Monteith, Voils, & Ashburn-Nardo, 2001). The fact that not every participant reaches awareness regarding the attitude being measured has been useful when comparing the performance of people who are and people who are not aware during testing. Awareness about the attitude being measured does not affect performance (Ashburn-Nardo, Voils, & Monteith, 2001; De Houwer, 2002).

Clinical Results

Regarding the IAT specifically, particularly in clinical research, results obtained by the IAT are inconsistent. Regarding specific phobias, the IAT has proven to have good results. In a study differentiating people with spider phobias and snake phobias. The IAT successfully discriminated between both groups, and classified correctly 92% of respondents when compared to a healthy control group (Teachman, Gregg, & Woody, 2001; Teachman & Woody, 2003). The IAT has been proven helpful to

distinguish between generally fearful participants from controls (Ellwart, Rinck, & Becker, 2006; Huijding & de Jong, 2007). However, these results were not replicated when, instead of images of spiders and snakes, experimenters used only spider and snake-related words (de Jong, van den Hout, Rietbroek, & Huijding, 2003). In this same group of patients, a baseline IAT predicted performance in a behavioural avoidance task, and significant changes in the IAT were observed after a 3-session exposure treatment, while it remained stable for a group who did not receive treatment. Changes remained stable in the IAT after a 2 months follow-up (Teachman & Woody, 2003).

Regarding social phobia, the IAT has been extensively used. It had been found to predict negative outcomes during social interactions (De Jong, Pasman, Kindt, & Van den Hout, 2001). This prediction also held for participants who did not have a diagnosis of social phobia. However, diagnosed participants showed significantly higher anxious self-associations (de Jong et al., 2003). These results have been replicated by various researchers (Gamer et al., 2008; Tanner, Stopa, & De Houwer, 2006). Egloff and Smuckle (2003) found that healthy participants trying to make a good impression for a job application could not manipulate their anxiety IAT, even when instructed about how to fake responses. According to theoretical expectations, a self-calmness IAT reflected changes obtained by a successful cognitive-behavioural intervention for socially phobic patients (Gamer et al., 2008).

Less consistent results were obtained by the IAT regarding panic disorder. While an IAT targeting associations between self and calmness was able to differentiate controls from disordered patients (Teachman, Smith-Janik, & Saporito, 2007), both groups scored equally in an IAT capturing associations between bodily changes and anxiety (Teachman, 2005). However, changes in the calmness IAT during psychotherapy predicted further changes in panic disorder severity (Teachman, Marker, & Smith-Janik, 2008).

Self-reported OCD symptoms are not related to IATs measuring associations between self and danger, self and immorality, and unwanted thoughts and their importance (Teachman & Clerkin, 2007; Teachman, Woody, & Magee, 2006). After a belief manipulation procedure, giving importance to intrusive thoughts and their moral significance, increased scores in the last IAT only.

In patients with PTSD both instrument- and clinician-rated severity were related to an IAT capturing the implicit association within self-and vulnerability (Engelhard,

Huijding, van den Hout, & de Jong, 2007). Note that the IAT's results did not predict the appearance of PTSD symptoms, but were a consequence of it.

Other disorders have been minimally studied with the IAT. Patients with body dysmorphic disorder showed lower implicit self-esteem as measured by the IAT, but this result was not replicated in a sample of body dysmorphic symptomatic students (Buhlmann, Teachman, Gerbershagen, Kikul, & Rief, 2008; Clerkin & Teachman, 2009). An IAT capturing implicit associations between sex and adults vs sex and children was able to correctly classify 78% of a sample of paedophilic individuals, but it had a great number (42%) of false-positives {Gray, 2005 #2178}. A modification of the IAT found increased anxiety associations to sexual stimuli in women with dyspareunia. (Melles et al., 2014). However, these results were not replicated in an affective Simon task (Brauer, de Jong, Huijding, Laan, & Ter Kuile, 2009).

In eating disorders, there are also mixed results that are difficult to interpret. For example, regarding overeating, an IAT measuring implicit attitudes towards food high in fats, found that both obese people and controls showed negative implicit attitudes towards these foods, with an increased negative attitude in obese people (Maison, Greenwald, & Bruin, 2001; Roefs & Jansen, 2002). These results run against expectations, nevertheless similar results were found in children (9-18 years old), with both obese and healthy children showing increased positive associations towards healthy food in contrast to unhealthy food, and an implicit preference for hobbies over palatable foods (Craeynest, Crombez, Haerens, & De Bourdeaudhuij, 2007). However, this same group of investigators found that lean children implicitly associate themselves with non-fatty food over fatty foods, while obese children did not show such difference (Craeynest, Crombez, De Houwer, Deforche, & De Bourdeaudhuij, 2006). A 12-week weight-loss treatment for these children had the expected results in weight reduction, but there were no changes in the IATs (Craeynest, Crombez, Deforche, Tanghe, & De Bourdeaudhuij, 2008). Adults who present retrained eating also show implicit preferences for non-fatty foods (Vartanian, Polivy, & Herman, 2004).

Of special interest for this thesis are results of the IAT in depressive disorders. They will be thoroughly reviewed in Chapter 4. But it is licit to advance that results are inconsistent, with depressed patients showing high implicit self-esteem in some studies. While some studies show some indexes of predictive validity for the IAT in depression, others find that diagnosis of depression is unrelated to IAT scores (Roefs et al., 2011).

In conclusion, it is apparent that in spite of the large amount of investigations utilising the IAT since its inception 15 years ago, results are still mixed and their interpretation remains debatable. On the other hand, although the use of the IAT is widespread in most topics within psychology, there are still some fields to cover. The next chapter develops a novel version of the IAT with the objective of measuring implicit attitudes towards parenting and their potential to predict attachment transmission, an unexplored territory in IAT investigations.

Subsequent chapters will utilise an adaptation of the Self-Esteem IAT in different samples. As we will see through the present thesis, measurement of implicit self-esteem is common in the literature, but it is perhaps the most difficult to interpret, and the topic that has been most questioned regarding its criterion validity (Buhrmester et al., 2011, p. 366). As this last authors posits: "care must be taken if the development of new research methods morphs into the introduction of new psychological constructs. For example, the mere presence of new labels such as *implicit self-esteem* tends to legitimise associated measures by implying that the construct has already been validated".

Chapter 2: Mothers' Implicit Attitudes towards Parenting: Relationships to Parenting Quality and Offspring's Attachment Style

Introduction

This chapter describes a study carried out with the objective of measuring the impact of mothers' implicit, non-conscious attitudes towards parenting and attachment on the attachment security of their young infants, through the use of the Implicit Association Test, which was originated within the context of social and cognitive psychology but has scarcely been used in the field of attachment research. Current measures of attitudes towards parenting and attachment suffer the same problems we have reviewed in the previous chapter, regarding their explicit nature and proneness to manipulation. Concurrently, there has been increased interest in the various mechanisms regulating intergenerational transmission of attachment (Belsky, 2005; Berthelot et al., 2014; Cyr, Dubois-Comtois, Pascuzzo, Béliveau, & Ellen, 2014; Tarabulsy et al., 2005). Given that much of what governs parenting behaviours is automatic and not readily available for introspection, it seems relevant to utilise implicit measures to learn more about the determinants of attachment transmission.

This chapter starts with a brief review of attachment theory and its importance in understanding parenting behaviour. Then, I revise the issues concerning assessment of non-conscious psychological constructs and their application to this particular theory. The paper then describes the present study, followed by a summary and discussion of results and suggestions for future research in this area.

Attachment Theory and Parenting Behaviour

Originally formulated by John Bowlby , attachment theory addresses the origins and mechanisms of people's characteristic ways of relating in intimate relationships to "attachment figures", often one's parents, children, and romantic partners. From birth, the interactions of an infant with his/her primary carers are thought to establish a base for personality development and to mould subsequent close relationships, expectations of social acceptance, and attitudes to rejection. When an infant's attachment figure (usually the mother) repeatedly provides stability and safety in moments of stress, a "secure base" is created (Bowlby, 1973, p. 182), which allows the infant to explore his/her surroundings. Thus, the child creates a set of mental models of him/herself and others in social interactions ("internal working models") based on repeated interactions with significant others. These early attachment relations are crucial for the acquisition

of capacities for affect and stress regulation, attentional control, mentalization, and for the infant's sense of self-agency (Lorenzini & Fonagy, 2013).

According to Bowlby (1969), human beings are born with an incipient capacity to provide protection and support to others who are dependent or in need. These tendencies, like empathy and compassion, are organised by an innate caregiving system, whose goal it is to reduce other people's suffering, protect them from harm, and foster their growth and development (Adam, Gunnar, & Tanaka, 2004). That is, the caregiving system is designed to provide a safe haven and a secure base for exploration and growth; therefore it is the complement of the attachment system (Green, Furrer, & McAllister, 2007). Following Collins et al. (in Mikulincer & Shaver, 2009), caregiving is activated when another person has to cope with danger, stress, or discomfort, when they are either seeking help or would benefit from it. This system is also activated when another person has an opportunity for exploration, learning, or mastery and either needs help in taking advantage of the opportunity or seems eager to talk about and be validated for his or her efforts and accomplishments. When this system is activated, the caregiver responds with behaviours aimed at relieving the needy person's distress, supporting his or her coping efforts, and/or providing a secure base for exploration, growth, and development.

It is thought that effective caregiving is organised around an empathic stance towards others, which includes what attachment theorists call sensitivity and responsiveness (Ainsworth, Blehar, Waters, & Wall, 1978; Wolff & Ijzendoorn, 1997). Sensitivity implies attunement to, and accurate interpretation of, another person's signals of distress, worry, or need, and responding in synchrony with the person's support-seeking behaviour. Responsiveness is helping the other feel loved, understood and cared-for, through validating and respecting the other's thoughts, beliefs needs, and feelings (Mikulincer & Shaver, 2009). A failure of these capacities for sensitivity and responsiveness can cause someone who seeks support to feel misunderstood, disrespected or burdensome. In childhood this leads to insecure attachment (Brown, Mangelsdorf, & Neff, 2012; Shah, Fonagy, & Strathearn, 2010; Wolff & Ijzendoorn, 1997); in adulthood it could lead to demoralisation and withdrawal from a relationship (Mikulincer & Shaver, 2009; Schoenmaker et al., 2015).

Further, the caregiving system is assumed to be a function of the caregiver's own attachment system. Many studies have indeed found that attachment is consistently "transmitted" to offspring (Shah et al., 2010; Van Ijzendoorn, 1995). Secure mothers

tend to raise secure children, while children of avoidant mothers tend to show an anxious attachment pattern (Shah et al., 2010). Infant attachment security has been associated to increased responsiveness to infant signals (Ainsworth et al., 1978; Belsky, Rovine, & Taylor, 1984; Isabella, Belsky, & von Eye, 1989). Moreover, it has been reported that infants' security of attachment increased after mothers completed caregiver sensitivity training (Boom, 1994; Howes, Galinsky, & Kontos, 1998; Moss et al., 2011).

It is noteworthy that sensitive responsiveness can be subsumed under the concept of mentalization, the explicit and implicit imaginative capacity to understand behaviour (in ourselves and others) as caused by mental states (Fonagy, Gergely, & Jurist, 2004). The mentalization model was first outlined in the context of a large empirical study in which security of infant attachment with each parent proved to be strongly predicted not only by that parents' security of attachment during the pregnancy, but even more so by the parents' capacity to understand their childhood relationships with their own parents in terms of states of mind (Fonagy, Steele, & Steele, 1991).

This is in line with Bowlby's (Bowlby) idea that individual differences in attachment arise as a result of the availability, responsiveness, and supportiveness of attachment figures. Interactions with figures who are available, sensitive, and responsive to one's bids for support facilitate the development of a sense of attachment security. In contrast, when attachment figures are not reliably available and supportive, a sense of security is not attained and negative internal working models are formed (e.g. worries about others' intentions, doubts about self-worth and self-efficacy). These working models eventually result in a trait-like attachment style (Dykas & Cassidy, 2011; Mikulincer & Shaver, 2012). This distorted attachment system is activated in moments of interpersonal distress. Childrearing is a challenging and –at times- stressful activity, thus it is expected that mothers with an insecure attachment style will react with a lack of responsiveness and sensitivity.

When asked to remember details of their own childhood, mothers with insecure attachment representations report more negative recollections of early parental caregiving, particularly rejection and the discouragement of independence. These same mothers, when their own infants were 2 months old, experienced heightened levels of maternal separation anxiety (Lutz & Hock, 1995). From very early in the empirical enquiry into infant attachment, various studies have shown that maternal anxiety appears as a relatively strong predictor of infant attachment security, through a negative

impact on maternal sensitivity (Del Carmen, Pedersen, Huffman, & Bryan, 1993; Mount, Crockenberg, Jó, & Wagar, 2010; Seymour, Giallo, Cooklin, & Dunning, 2014; Shin, Park, Ryu, & Seomun, 2008; Stevenson-Hinde, Chicot, Shouldice, & Hinde, 2013; Stevenson-Hinde, Shouldice, & Chicot, 2011).

More recently, a number of researchers have explored the multiplicity of attachment models available to people. That is, although people may be guided by one global attachment orientation, research reveals that they actually have multiple attachment models available in memory (Bartz & Lydon, 2004; Shaver & Mikulincer, 2002b). This finding, together with findings concerning the increased genetic contribution to attachment during adolescence, may partly explain the differences in attachment security between siblings (Fearon, Shmueli- Goetz, Viding, Fonagy, & Plomin, 2014; Van Ijzendoorn et al., 2000). These considerations make the prediction of attachment styles a challenging enterprise, given the multiple possible causal factors and the relationship between them, which in turn imply the development of various strategies to measure these causal factors.

Assessment Issues in Attachment Theory

In general psychological science, it has long been believed that relatively unconscious processes affect the way that parents respond to and interact with their young children. In the field of attachment, these unconscious processes correspond to the "internal working models" (IWMs) (Bowlby, 1969 p.80). Because some features of IWMs are believed to work outside of awareness (Shaver & Mikulincer, 2002a), they are difficult to study. In spite of their importance, little is known about the role played by unconscious thoughts, feelings and attitudes in parenting and attachment transmission (Bartholomew & Moretti, 2002; Dewitte et al., 2008; Shin et al., 2008).

However, the influence of Bowlby's attachment theory in social and cognitive psychology has led to the operationalisation of internal working models as cognitive-affective relationship schemas. The elements of a relational schema include an interpersonal script for the interaction pattern, a self-schema for how the self is experienced in that interpersonal situation, and a schema for the other person in the interaction (Baldwin, 1992; Maier, Bernier, Pekrun, Zimmermann, & Grossmann, 2004). This is in line with Bowlby's original idea that the activation of one internal working model (e.g. the activation of the caregiver's model) gives rise to the activation of related models of self and others in the social environment (Bowlby, 1973).

As a consequence of this convergence point, which constitutes a common language between psychodynamic and cognitive domains (Blatt, 2011; Kihlstrom, 1987; Kihlstrom, 1994; Kihlstrom et al., 1992; Kihlstrom et al., 2000), attachment researchers have started to rely on experimental social-cognitive methods for investigating the accessibility and organisation of attachment representations using implicit measures. These methodological advances appear well suited for measuring unconscious attitudes. As pointed out in the previous chapter, implicit measures can be defined as measures of psychological processes that are uncontrolled, unintentional, goal independent, purely stimulus driven, autonomous, and/or unconscious (De Houwer et al., 2009). They attempt to measure the comparative strength of evaluations and preferences that are automatically activated and exist outside of conscious awareness or conscious control (Greenwald et al., 2002; Posner & Snyder in Maier et al., 2004). Such measures are the "opposite" of explicit measures, i.e. self-report questionnaires, which are more vulnerable to conscious control and social desirability (Nosek et al., 2005).

Various types of implicit measures have been used in the study of attachment, chiefly of adult attachment. It has consistently been found that adult attachment styles, usually measured by self-report are related to implicit and automatic measures of attachment, particularly in studies using priming techniques (Gillath et al., 2006; Maier et al., 2004; Mikulincer, Birnbaum, Woddis, & Nachmias, 2000; Mikulincer, Gillath, & Shaver, 2002; Mikulincer, Hirschberger, Nachmias, & Gillath, 2001; Mikulincer, Shaver, Sapir-Lavid, & Avihou-Kanza, 2009). However, there is always a gap between conscious and unconscious models of attachment: there is evidence, coming from priming studies, that the higher an individual's security score, the more congruent his or her conscious and unconscious representations of self and others are. For example, it has been found that subliminal priming of the sentence "my mom rejects me" before answering questions regarding the relationship of the self with parental attachment figures, was unrelated to an attachment self-report (the Inventory of Parent and Peer Attachment, IPPA), and partially related to attachment style measured by the Adult Attachment Interview (AAI). Specifically, higher attachment security in the IPPA predicted extended reaction times in the priming condition to questions assessing security, which implies that positive answers in self-report are a defensive construction (an idealisation of the attachment relationships), which is fragile to the threat posed by priming of maternal rejection. Thus, while self-reported attachment fails at accounting for defensive and self-presentation biases of dismissive individuals (yielding a highly

secure attachment profile), both the AAI and the priming model showed unbiased profiles, which resulted to be dismissive (Maier et al., 2004).

The Implicit Association Test in Attachment Research

In this study two types of Implicit Association Tests (IATs) were used to assess automatic attitudes towards parenting and attachment. The IAT is a flexible, computer-based test that requires the participant to categorise two groups of words and/or pictures, with one group reflecting a relevant theme (e.g. ethnicity: pictures of white and black people) and another a value dimension (e.g. good versus bad). The IAT can detect the strength of the association of one of the themes with one of the valences relative to the other theme with the other valence, depending on the response-time. For example, people with an implicit preference for white people will show decreased reaction times when shown pictures of white people and are required to press the same key associated with good words, as both notions (white and good) are assumed to activate the same network of associations. This instrument has proven to be a good predictor of behaviour relevant to the specific attitude in question and, to a lesser extent, of directly expressed views on the same topic (Greenwald & Farnham, 2000; Greenwald et al., 1998; Greenwald et al., 2003; Greenwald et al., 2009; Nosek et al., 2005, 2007a; Rudman, 2008). For a more detailed description of the IAT, see Chapter 1.

The IAT has barely been used in the field of attachment research, but the results are encouraging. Zayas and Shoda (2004 , cited with permission; 2005) used an IAT to measure automatic evaluative associations regarding the self and a significant other. Their study revealed that scores on a Partner IAT (using the categories *partner, not-partner* and *pleasant, unpleasant)* were related to explicit measures of adult attachment styles, namely the Relationship Questionnaire (RQ) and the Experiences in Close Relationships Questionnaire (ERC), and to other relationship outcomes such as relationship satisfaction, commitment, and relationship length. Individuals that were securely attached to their partners showed stronger positive evaluations, while avoidant individuals showed weaker implicit positive evaluations of their partner. There was no clear relationship between implicit evaluations and anxious attachment. A Self-IAT (using the categories *me, not-me* and *pleasant, unpleasant)* was not related to attachment orientation. It could be argued that the Partner-IAT functions more as a test of preference rather than attachment, however an IAT measuring implicit evaluations of the subjects' mothers predicted attachment style with a current romantic partner,

following the same trends evidenced by the Partner IAT. None of these IAT results were associated with questionnaire measures assessing explicit attachment towards their mothers and partners. These results seem to support the idea that the most enduring attachment representations are in turn, the most automatic and therefore the least conscious (Bowlby, 1969).

Another research team (Dewitte et al., 2008) used IATs that have shown validity in assessing implicit relational self-esteem (the sense of personal self-worth achieved through relationships with significant others) and implicit anxiety, and found that they were meaningfully related to individual differences in attachment style as measured by self-report questionnaires. They also found that these IATs predicted attachment-related thoughts and feelings beyond self-reported attachment.

With regards to attachment research, the IAT is normally used in pairs; that is to say, two IATs are administered to each subject in the sample. In general, one IAT uses contrasting categories "self vs. other" and "good vs. bad". The other IAT uses "attachment figure vs. other" and "available vs. unavailable". These categories are conventionally used because they seem to be congruent with attachment theory, in the sense that positive and negative models of the self, and the sense of availability or absence of the other, combine to form different attachment styles as show in Table 1 (Bartholomew & Horowitz, 1991; Veletanlic, 2007). These categories have shown different convergent and divergent validity depending on the target attachment figure. In a sample of romantically involved adults, Veletanlic showed that implicit associations of "self" with "good" were related to implicit self-esteem, and implicit associations of "partner" with "good" were related to implicit sociability. However, these implicit effects were unrelated to measures of explicit attachment, measured by the ECR and the Situational Attachment Questionnaire. Conversely, these same categories, when measuring adult attachment to mothers instead of romantic partners in a Chinese sample of undergraduates, were positively related to measures of explicit attachment (ECR and RQ). In general, it seems that the various IATs used in attachment research converge with explicit measures of the same constructs, but predict attachment

		Positive	Negative
MODEL OF	Positive	Secure	Preoccupied
OTHERS	Negative	Dismissing	Fearful

Table 1: Theoretical model of attachment. Adapted from (Bartholomew & Horowitz, 1991)

beyond explicit measures and are less affected by social desirability and self-presentation biases.

The use of the IAT in attachment research has hitherto been limited to the study of partner attachment and adult attachment styles. The present study aims to use this methodology to assess whether theoretically expected implicit attitudes can be detected that are linked with the quality of parenting and the security of infant-parent attachment.

Hypotheses

Mothers who are more sensitive and responsive to their infants (as assessed by observing a structured play interaction between a parent and their infant) and who have infants who are securely attached (as assessed in Ainsworth's Strange Situation procedure) are expected to implicitly value attachment more strongly. In order to measure implicit attitudes towards attachment, we developed two new IATs measuring two aspects of attachment. In the first, participants categorised pictures depicting scenes of parents providing comfort to infants versus pictures showing an unaccompanied adult in leisure situations. Participants also had to categorise words belonging to the categories GOOD and BAD (see Table 3 and Figure 2).

In a second IAT participants categorised the same set of value items (good versus bad) and a set of pictures either showing their own infant or an unfamiliar infant.

It is expected that mothers of securely attached infants will show significantly more positive implicit attitudes towards parental activities and attachment issues than mothers of insecure infants, as measured by two IATs.

Given that most studies using implicit measures in attachment find that implicit attitudes are related to both explicit attitudes and behaviour, it is also expected that the relationship between implicit attitudes and attachment security will be mediated by the quality of observed parenting behaviour.

Likewise, it is expected that the relationship between implicit attitudes towards parenting and attachment security will be mediated by explicit self-report measures of parenting.

It is also expected that maternal levels of self-reported psychiatric symptomatology, particularly anxiety, will mediate the relationship between implicit attitudes towards attachment and parenting, and offspring's attachment security.

Finally, we expect that the relationship between implicit associations towards attachment and parenting, and infants' attachment security will be mediated by mothers' mentalization abilities.

In general, this study has the potential to significantly advance our understanding of the processes that give rise to differences in the quality of parenting that children receive, and of the mechanisms underlying attachment transmission.

Within the context of this thesis, this study represents an initial approximation to the use and analysis of the IAT and its potential for addressing individual differences in parenting attitudes and behaviour, and in offspring's attachment status. The results of this study will warrant (or not) a more in-depth analysis of the relationship of implicit self-esteem and attachment. As mentioned at the beginning of this thesis, its goal is to assess the usefulness and pertinence of the IAT for the study of various constructs in social cognition. It is then important to address the phenomenon of attachment as one of the central concepts in social-cognitive science.

Methods

Design

The present study follows a cross-sectional correlational design. Measures of attachment and parenting quality were collected within a timespan of no longer than two months. The measures were carried out during the years 2007 and 2008 at the Developmental Laboratory of the Anna Freud Centre in London, UK, within the context of a larger study on the validation of the Reflective Function Questionnaire (Fonagy, Luyten, & Perkins, 2015).

Procedure

Specifically, mothers and their infants were observed at the laboratory at 10 months of age in order to measure the quality of parent-child interaction. These observations took between 30 and 40 minutes. During this same visit, mothers were asked to complete a small battery of questionnaires and two computerised IAT tests. Two months later, mothers and infants were asked to return to the Centre in order to assess the infant's attachment security using Ainsworth's Strange Situation observation procedure (details provided below). At that point, mothers were given the Reflective

Function Questionnaire and the Parental Reflective Function Questionnaire, to be returned to the centre by post.

Participants

The sample was drawn from a large database of parents from the community in Northwest London who had voluntarily agreed to be contacted regarding developmental studies. Volunteers on these databases were initially recruited by researchers visiting groups that parents and their infants would attend, such as playgroups, baby massage groups, libraries, baby health clinics and breastfeeding support groups. Researchers approached parents, explaining to them about the various research projects carried out at the Developmental Laboratory of the Anna Freud Centre. Other participants registered on the database after seeing posters and leaflets about the work of the Anna Freud Centre left in places parents would visit.

The total sample was composed of 124 mother-infant dyads. They all completed the full assessment battery with the exception of the RFQ and PRFQ, which were returned to the Laboratory via post by a subsample of 64 mothers (51.6% of the total sample).

Table 2 shows the demographic composition of the general sample and that of the subsample.

Table 2also shows the distribution of attachment security and attachment style. Both in the subsample which returned the RFQ and PRFQ and in the general sample, the distribution of attachment styles shows an higher percentage of securely attached infants and a lower presence of both avoidant and anxious styles in comparison with other general population samples, expected to be 63%, 16%, and 21% respectively (Main, Kaplan, & Cassidy, 1985). There were no infants classified as disorganised. The subsample that completed the RFQ and PRFQ did not show any significant differences when compared to the total sample, thus subsequent analyses were carried out in the total sample.

	Participants Who Completed RFQ and PRFQ (n = 64)	Total Sample (n = 124)
Baby's Gender		
Female	35 (54.7%)	67 (54.0%)
Male	29 (45.3%)	57 (46.0%)
Mother's Age Mean in years (SD)	35.05 (3.46)	35.07 (4.40)
Father's Age Mean in years (SD)	37.72 (4.66)	37.11 (5.20)
Number of Baby's Siblings		
0	22 (34%)	47 (37.9%)
1	16 (25.0%)	29 (23.4%)
2	2 (3.1%)	7 (5.6%)
3	2 (3.1%)	3 (2.4%)
No Information	22 (34.4%)	38 (30.6%)
Mother's Level of Education		
GCSEs or Equivalent	3 (4.7%)	5 (4.0%)
A-Levels	4 (6.3%)	8 (6.5%)
NVQ	1 (1.6%)	6 (4.8%)
HND	9 (14.1%)	14 (11.3%)
BA/BSc	25 (39.1%)	57 (46.0%)
Masters/PhD	22 (34.4%)	34 (27.4%)
Father's Level of Education		
GCSEs or Equivalent	5 (7.8%)	12 (9.7%)
A-Levels	3 (4.7%)	7 (5.6%)
NVQ	Nil	1 (0.8%)
HND	9 (14.1%)	15 (12.1%)
BA/BSc	26 (40.6%)	48 (37.7%)
Masters/PhD	17 (26.6%)	32 (25.8%)
No information	4 (6.3%)	7 (5.6%)

Table 2: Demographic description of the Reflective Function subsample and the full sample

	Participants Who Completed RFQ and PRFQ (n = 64)	Total Sample (n = 124)
Marital Status		
Married and Co-habiting	44 (68.8)	79 (63.7%)
Married and Living Apart	Nil	2 (1.6%)
Unmarried and Co-habiting	13 (20.3%)	32 (25.8%)
Single	7 (10.9)	11 (8.9%)
Household income per annum		
Less than £10,000	2 (3.2%)	7 (5.9%)
£10,000 – 20,000	1 (1.6%)	3 (2.5%)
£20,000 – 30,000	9 (14.4%)	15 (12.7%)
£30,000 – 40,000	5 (8.1%)	11 (9.3%)
£ 40,000 – 50,000	15 (24.2%)	24 (20.3%)
More than £50,000	30 (48.4%)	58 (49.2%)
Baby's Attachment Security		
Secure	47 (73.4%)	90 (73.2%)
Insecure	17 (26.6%)	33 (26.8%)
Three-way Attachment Classification		
Avoidant (A)	9 (14.1%)	20 (16.3%)
Secure (B)	47 (73.4%)	89 (72.4%)
Ambivalent (C)	8 (12.5%)	14 (11.4%)
IAT Comfort-Leisure Mean D score (SD)	76.69 (231.03)	77.36 (228.27)
IAT MyBaby Mean D score (SD)	304.93 (277.21)	327.28 (299.92)

Table 2: cont.

Measures

Observation of parent-child interaction. Procedures based on those described by Pederson and colleagues (Pederson & Moran) were used to measure key dimensions of parenting by observing mothers and infants in several naturalistic situations. Mothers were instructed to interact with their infant as they normally would. Interactions were recorded with a digital video camera. In one 5-minute segment mothers were asked to play freely with their infants in the absence of any toys. In a second segment, mothers played with their infant with toys. In a third, mothers were asked to interact normally with their infant while carrying out a distracting task (so that the parent has to divide her attention between the baby and the task, namely filling out a questionnaire). In a fourth scenario, mothers needed to support their baby in completing an age-appropriate developmental assessment.

Interactions at each time-point were coded with the Coding Interactive Behaviour manual (CIB) (Feldman, 1998). The CIB coding system is a global measure that looks at parent, child and dyadic affective states and interactive styles. This measure is typically used with adults and children aged between 2 and 36 months. The CIB is broken down into 43 codes that are rated on 5-point Likert scales. There are 21 parent codes, 16 child codes and 5 dyadic codes. All 43 codes can be calculated into subscales, consisting of parental sensitivity (CIB-S, Cronbach's alpha = .92; Ferber & Feldman, 2005), intrusiveness (CIB-I, α = .78; Ferber & Feldman, 2005) and limit setting (CIB-L, α = .86; Keren, Feldman, & Tyano, 2001), child involvement (CIB-CI, α = .86; Feldman, 2000), withdrawal and compliance (CIB-W, α = .86; Keren et al., 2001), and dyadic reciprocity (CIB-R, α = .89; Ferber & Feldman, 2005) and negative states (CIB-N, α = .71; Feldman, Weller, Sirota, & Eidelman, 2003).

The coding system has shown good psychometric properties across various cultures and it is sensible to differences stemming from infant age, interactive partner, cultural background, biological and social-emotional risk conditions, and change following intervention (Bartling, Wiefel, Klapp, Jonkman, & Lenz, 2006; Feldman, 2000; Feldman, Greenbaum, Mayes, & Erlich, 1997; Feldman, Keren, Gross-Rozval, & Tyano, 2004) (Ainsworth et al., 1978) (Meins, 1997; Meins et al., 2002) (Isabella, 1993; Keren, Dollberg, Koster, Danino, & Feldman, 2010; Lohaus, Keller, Ball, Elben, & Voelker, 2001). All interactions were coded by two trained researchers.

Implicit attitudes toward attachment. The Implicit Association Test (IAT) is a computerised test that requires participants to categorise two groups of stimuli (pictures and words) along pre-defined lines. For each stimulus appearing in the centre of the screen, the time taken to make a categorisation using one of two keys is recorded in milliseconds by the computer. This test consists of five steps (Figure 2). Step (a): the experiment begins with a participant categorizing the target concept stimuli (i.e. ATTACHMENT versus NON-ATTACHMENT pictures, or photos of the participant's own baby versus photos of unfamiliar infants) as quickly as they can, using two different keyboard keys. Step (b): This is a second block in which the attribute words (GOOD versus BAD) are categorised using the same keys as in the previous step. Step (c): In the third block the concept and attribute stimuli are combined and participants must categorise them using the keys as before. This time, however, on one side of the screen the concept and the attribute will be paired: ATTACHMENT with BAD, or with GOOD, depending on the counterbalancing (see Step e). When ATTACHMENT is paired with GOOD, we call this the *congruent condition.* Step (d): In the fourth step, the concept items are categorised again, but this time using the opposite keys to those used previously. Step (e): Finally, in a fifth block the categories are combined again and the concept items must be categorised using the keys from the fourth block. The categories will be combined in the opposite way they were in step c. If ATTACHMENT is paired with BAD, we call this the *incongruent condition.* Given that the assumption underlying the IAT is that people will take more time in classifying concepts and attributions when they are less strongly associated (or incongruent), implicit attitude strength is thus determined by the difference in response time between steps c and e (in which the concept items and attribute items are paired in opposite ways). This difference between the incongruent and congruent conditions is called the *D* score (Greenwald et al., 2003). This is to say, when ATTACHMENT is more strongly associated with GOOD, then response times when this concept and attribute are paired will be shorter than the time taken to classify words and photos when ATTACHMENT is paired with BAD.

ATTACHMENT VS. NONATTACHMENT IAT

MY BABY VS. OTHER BABY IAT

Step a
BABY — LEISURE | OTHERS — MY BABY

Step b
GOOD — BAD / Joyful | BAD — GOOD / Greedy

Step c
BABY GOOD — LEISURE BAD | OTHERS BAD — MY BABY GOOD

Step d
BAD — GOOD / Gratifying | GOOD — BAD / Angry

Step e
BABY BAD — LEISURE GOOD | OTHERS GOOD — MY BABY BAD

Figure 2:The five steps of the IATs used in this chapter. Attachment/Leirure left, MyBaby IAT on the righ

Trial blocks. Both IATs were administered in trial blocks of 50 trials each. Each trial block started with instructions that described the category discrimination(s) for the block and the assignment of response keys (E or I) to categories. Reminder labels, in the form of category names were appropriately positioned to the left or right, and remained on screen during each block. Each new category discrimination, steps (a), (b), and (e) described above, consisted of a practice block of 50 trials followed by a block for which data was analysed. IAT effect is usually biased toward indicating greater strength of associative pairings involved in the first of the two combined tasks, namely step c (Greenwald et al., 1998; Greenwald et al., 2003). That is to say, participants who first sort ATTACHED with BAD and NON-ATTACHED with GOOD and then sort the reverse configuration are likely to show a stronger indication of implicit preference for NON-ATTACHED over ATTACHED than participants who first sorted ATTACHED with GOOD and NON-ATTACHED with BAD. This means that the order of the blocks had to be counter-balanced between participants.

Timing Details. The first trial started 1.5s after the reminder display appeared. Verbal stimuli were presented in black letters against a light grey background screen, vertically and horizontally centred in the display. Photographic stimuli were displayed in the same location and subtended a similar visual angle to the participant's eye. Stimuli remained on screen until the subject's response. The subject's key press response initiates a delay (inter-trial interval, where no stimulus is presented) before the next trial's stimulus. For all simple categorisation and combined-tasks, the inter-trial interval was 400ms. Throughout the experiment, after any incorrect response, the word *error* would immediately replace the stimulus for 30ms, lengthening the inter-trial interval by 300ms.

Stimuli. Words and pictures were selected randomly from a list previously categorised by expert judges as attachment or non-attachment related. These words were categorised as GOOD or BAD (Table 3). Pictures of the participant's child were taken using a digital camera on the day of testing. Each picture was paired with photos of infants of similar age and gender. Once a stimulus was drawn from the stimulus pool it was not replaced (independently for each subject) until the available stimuli for a task were exhausted, at which point the stimulus pool was replaced if more trials were needed. In all combined tasks, items for the target-concept discrimination and the attribute discrimination appeared on alternating trials.

GOOD	BAD
Joyful	Uncomfortable
Pleasurable	Disappointing
Comforting	Awkward
Loving	Tedious
Fond	Stressful
Fulfilling	Angry
Enjoyable	Greedy
Gratifying	Irritating

Table 3: GOOD and BAD word stimuli for both IATs

Explicit attitudes toward parenting and attachment. *Parental Modernity (PM) Scale of Child-rearing and Education Beliefs (Ideas about Raising Children).*

Mothers completed this thirty-item, Likert-type questionnaire during the first visit. The instrument was designed to measure traditional authoritarian and progressive democratic beliefs of parents (Schaefer & Edgerton, 1985). The scale yields a total score and two sub-scores: Progressive Beliefs (reflects attitudes favouring self-directed child behaviour) and Traditional Beliefs (reflects attitudes that child behaviour should follow adult directives) (Campbell, Goldstein, Schaefer, & Ramey, 1991). Initial studies with the parental modernity measures revealed split-half reliability to be 0.90 and test-retest reliability to be 0.84 (Holden & Edwards, 1989).

Mother's Object Relations Scales (MORS). The 44-item Mother's Object Relations Scales (MORS) instrument was developed as a means of quantitatively assessing core features of mothers' ideas and beliefs about infants particularly in relation to attachment. Principal components analyses of datasets collected in Hungary and Great Britain have shown common latent structures in mothers' responses to the instrument, supporting a prediction of two underlying axes of mothers' perception of infants' "warmth-coldness" and "withdrawal". The instrument has also been found to correlate significantly with ratings of infant temperament and maternal mental states (Oates & Gervai, 2000).

Maternal Separation Anxiety Scale MSAS (Parental Care). The 21 items on this questionnaire constitute the Maternal Separation Anxiety factor from the Maternal Separation Anxiety Scale (Hock, McBride, & Gnezda, 1989). Each item is a 5-point Likert scale ranging from "strongly disagree" (1) to "strongly agree" (5). This questionnaire assesses the mother's "level of worry, sadness, and guilt when separated from her infant; and most importantly for the current purposes her beliefs about the importance of exclusive maternal care; her beliefs that her child prefers her care and is better off in her care; and her beliefs about her child's abilities to adapt to non-maternal

care" (DeMeis, Hock, & McBride, 1986, p. 628). The internal consistency (Cronbach's alpha) for this factor was 0.90 in its validation studies (Hock et al., 1989).

Parenting Stress Index (PSI) (Feelings about Parenting, FAP). Mothers also completed a 30-item, modified version of the 101-item PSI, which was used and validated in the large scale NICHD Study of Early Child Care (Early Child Care Research Network NICHD, 1994). The PSI is designed to identify parent-child systems that are under stress and at risk for development of dysfunctional parenting, behaviours or behaviour problems in the child involved. While the scale includes some items unrelated to parenting attitudes, several dimensions of the questionnaire address parenting attitudes, particularly to the more difficult aspects of childrearing. Cronbach's alphas for the 3 subscales of the Parental Characteristics domain of the PSI have been found to be Attachment = 0.52; Restrictions of Role = 0.75; Sense of Competence = 0.73; and the overall scale = 0.82 (Brown, 2011).

Reflective Function Questionnaire (RFQ-46). The term reflective function or mentalization refers to the capacity to reflect on internal mental states such as feelings, wishes, goals, and attitudes, both with regard to the self and with regard to others. Studies suggest that this capacity first develops in the context of secure attachment relationships. This 46-item questionnaire assesses the reflective function in adults with regards to the certainty and uncertainty with which individuals address mental states. It yields two sub-scores: Certainty of Mental States (RFQ-C) and Uncertainty of Mental States (RFQ-U). Initial validation studies show good psychometric properties (Cronbach's α = 8.4 and 8.1, respectively) (Luyten, Mayes, Nijssens, & Fonagy, 2013). Mothers completed this questionnaire in their homes after the second visit and posted it back to the laboratory, together with the Parental Reflective Function Questionnaire.

Parental Reflective Function Questionnaire (PRFQ). This 39-item test assesses parental reflective functioning or mentalizing, that is, the capacity to treat their own infant as a psychological agent. It yields three sub-scores with good psychometric properties: Pre-Mentalising Modes (PRFQ-PM), which measures a non-mentalising stance, malevolent attributions to the baby, and an inability to enter the subjective world of the infant (items include statements like: "My baby cries around strangers to embarrass me"). The second subscale, named Certainty About Mental States (PRFQ-CSM) reflects the tendency not to recognise the opacity of mental states (e.g., "I always know what my child wants"). The third factor measures Interest and Curiosity about mental states (PRFQ-IC) (e.g. "I like to think about the reasons behind the way my

child behaves and feels). Preliminary validation studies show evidence of reliability and validity for this measure (Luyten et al., 2013).

Infants' attachment. Ainsworth's Strange Situation Procedure (SSP). This is a widely used and well-validated procedure for assessing the security of the infant-parent attachment relationship (Ainsworth et al., 1978). The procedure involves two brief episodes in which the infant is separated from the parent in an unfamiliar room. In the first separation the infant is left in the presence of a stranger and in the second the infant is left alone in the room. These separation episodes last up to 3 minutes and are curtailed if the infant shows signs of considerable distress. The procedure lasts approximately 20 minutes although it is often less if the separation episodes are curtailed. Infant–mother dyads are assigned to one of four classifications: secure, avoidant, resistant, or disorganised. Upon reunion, infants classified as secure (B) seek out whatever contact is needed and calm easily in the presence of their mothers. Infants classified as avoidant (A) turn away from their mothers when distressed. Infants classified as resistant (C) show angry resistance to attempts by their mothers to calm them. Infants classified as disorganised (D) appear to lack a strategy for dealing with distress in the mother's presence or show a breakdown in their strategies for dealing with distress. All infants were assigned one of the four categories described above as well as a best-fitting three-way classification. The procedure was coded by trained coders, who have demonstrated reliability with international standards.

Mediation measures. In order to assess the possible indirect effects of a range of other factors, several additional measures were taken:

The Brief Symptom Inventory (BSI): A 53-item self-report designed to assess the psychological symptom status of psychiatric and medical patients, as well as individuals who are not patients. The BSI has been used in a wide variety of settings and applications. Alpha coefficients range from 0.71 to 0.89 for its different scales (Boulet & Boss, 1991; Derogatis & Melisaratos, 1983). All mothers in the study will be asked to complete this instrument at the 10-month visit.

Demographic information: An ad-hoc form was used to record parent's highest level of educational attainment, their self-defined ethnicity, their occupation and the occupation of the child's father (if known), the infant's birth weight and sex, and whether the infant has any siblings.

Infant temperament observed by mother: The Infant Behaviour Questionnaire - Short Version (IBQ-S) was filled out by mothers at the 10-month assessment point. This

measure is a well-validated 91-item questionnaire concerning a range of dimensions of infant behaviour, indicative of temperament, which is used widely in developmental research (Heilbig, Putnam, Gartstein, & Rothbart, 2010). The subscales utilised in this study were purely behavioural: Activity Level (internal consistency $\alpha = 0.70$), Distress and Latency to Approach Novel Stimuli ($\alpha = 0.80$), Soothability ($\alpha = 0.77$), Duration of Orienting ($\alpha = 0.76$), and Smiling and Laughing ($\alpha = 0.79$; Putnam, Helbig, Gartstein, Rothbart, & Leerkes, 2014).

Data Analysis

The IAT data was analysed using IBM SPSS 21 for Mac. In order to assess if differences in parents' implicit attitudes towards patenting and attachment issues are associated to the attachment style of their infants, ANOVA and T-tests were utilised. Subsequently, continuous measures of implicit attachment-related attitudes were extracted and entered as independent variables in a logistic regression predicting 12-month attachment security, mediated by explicit observed and self-reported parenting, maternal psychopathology, and reflective function. Mediation analyses were carried out using *PROCESS*, a plug-in macro for SPSS (Hayes, 2013).

Results

This was a healthy sample, with only 3 mothers presenting BSI scores above the clinical cut-off score. Demographic characteristics of the sample were not significantly associated to either IAT score or infant's attachment style.

D scores were analysed using ANOVA and independent samples T-tests in order to establish whether there are demographic group differences regarding implicit measures of attachment. These analyses yielded no significant results when the sample was grouped according to baby's gender, if the baby was first-born or had more siblings, household income, psychiatric symptomatology, parents' level of education, and mother's marital status.

Disconfirming our hypothesis, there were no statistically significant differences in the IAT D scores between mothers of securely and insecurely attached babies (Attachment/Leisure: $t = -0.011$, $p = 0.991$; MyBaby: $t = -0.826$, $p = 0.411$). There were no significant differences in D scores between mothers of babies with different attachment styles, classified by both the 3-way and 4-way methods (Table 5 and Table 4). These results might be reflecting the lack of power of the study. At the 0.05 significance level, group sizes in both 3-way and 4-way classifications would need group sample sizes of at least 44 and 48 participants respectively, to have enough power (0.80) to find large effect sizes (Cohen, 1992).

Measure	Source	SS	df	Mean Square	F	p
D IAT Attachment/Leisure	Between Groups	12329.97	2	6164.986	.113	.893
	Within Groups	6550055.38	120	54583.795		
	Total	6562385.35	122			
D IAT MyBaby	Between Groups	66835.37	2	33417.685	.432	.650
	Within Groups	9209123.97	119	77387.596		
	Total	9275959.34	121			

Table 4: One-Way analysis of variance of IAT D scores by 4-way attachment classification

Measure	Source	SS	df	Mean Square	F	p.
D IAT Attachment/Leisure	Between Groups	14740.61	2	7370.31	.135	.874
	Within Groups	6547644.74	120	54563.71		
	Total	6562385.35	122			
D IAT MyBaby	Between Groups	187684.30	2	93842.15	1.229	.296
	Within Groups	9088275.03	119	76372.06		
	Total	9275959.34	121			

Table 5: One-Way analysis of variance of IAT D scores by 3-way attachment classification

	1	2	3	4	5	6	7	8	9
1. *D* IAT Attachment/Leisure	1								
2. *D* IAT MyBaby	.133	1							
3. Parent Sensitivity	.031	.001	1						
4. Parent Intrusiveness	.164	.219	-.502	1*					
5. Parent limit setting	.058	-.054	.839	-.613	1				
6. Child involvement	.015	.008	.571	-.317	.521	1			
7. Child withdrawal	-.012	-.061	-.458	.257**	-.413**	-.651**	1**		
8. Dyad reciprocity	.018	-.017	.867	-.556	.808	.659	-.563	1	
9. Dyad negative states	-.006	.104	-.731	.549	-.738	-.552	.530	-.809	1

*Correlation is significant at the 0.05 level (2-tailed).

**Correlation is significant at the 0.01 level (2-tailed).

Table 6: Pearson correlations for IAT D scores and CIB subscales

Bivariate Correlations Between Mothers Implicit and Explicit Attitudes Towards Parenting and Control Measures: Correlations between IAT *D* scores and the various subscales showed a general lack of relationship between the IAT scores and explicit measures of patenting attitudes as assessed by the PM, MORS, MSAS, and PSI. Likewise, IAT *D* scores showed no relationship with control measures with a few exceptions. Moreover, these correlations were contrary to our hypothesis. For instance, an implicit preference for parenting and attachment was positively correlated with measures of distress and psychopathology[4]. The few variables that present significant correlations with the IAT *D* scores do not correlate with each other, with the exception of BSI measures of anxiety and psychoticism, which were strongly correlated: $r(105) = 0.69$, $p < .001$.

---

4 To test whether this strong correlation implied that both psychoticism and anxiety were directly influencing the D scores independently, a stepwise multiple regression was run with the D score as outcome, and psychoticism and anxiety as predictors in that order. Once psychoticism was controlled, anxiety ceased to be significant ($R2 = 0.076$, $F(1,104) = 8.614$, $p = 0.004$).

	1	2	3	4	5	6	7	8	9	10	11	12	BSI total
D IAT Attachment/ Leisure	1	.133	-.090	.028	-.029	-.029	.139	.005	-.117	-.044	.080	.033	.001
D IAT MyBaby	.133	1	-.037	-.002	.076	.131	.239	.082	-.013	.105	.277	.118	.118
Somatisation	-.090	-.037	1	.233	.242	.244	.225	.173	.189	.392	.279	.438	.466
Obsession-compulsion	.028	-.002	.233	1	.567	.550	.536*	.375	.363	.427	.507**	.762	.787
Interpersonal sensitivity	-.029	.076	.242	.567	1	.659	.591	.445	.426	.681	.612	.794	.813
Depression	-.029	.131	.244	.550	.659	1	.652	.442	.335	.590	.713	.798	.818
Anxiety	.139	.239	.225	.536*	.591*	.652*	1*	.457	.284	.555**	.696**	.786**	.781**
Hostility	.005	.082	.173	.375	.445	.442	.457	1	.250	.541	.281	.623	.619
Phobic anxiety	-.117	-.013	.189	.363	.426	.335	.284	.250	1	.369	.281	.391	.484
Paranoid ideation	-.044	.105	.392*	.427	.681**	.590**	.555**	.541**	.369**	1**	.551**	.795**	.781**
Psychoticism	.080	.277	.279	.507	.612	.713	.696	.281	.281	.551	1	.738	.748
General severity	.033	.118	.438	.762	.794	.798	.786	.623	.391	.795	.738	1	1.000
BSI TOTAL	.001	.118	.466*	.787**	.813	.818**	.781**	.619**	.484**	.781**	.748**	1.000**	1**

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Table 7: Pearson correlations for IAT D scores and BSI subscales

	1	2	3	4	5	6	7	8	9
D IAT Attachment/Leisure	1								
D IAT MyBaby	.133	1							
MORS warmth-coldness	.054	-.057	1						
MORS Invasive-withdrawal	-.092	.061	-.246	1					
PM progressive beliefs	.053	-.144	.131	-.009	1				
PM traditional Beliefs	.064	.095	-.074	.060	-.082	1			
PM Total	.058	.127	-.088	.080[**]	-.226	.971	1		
Parental Stress, FAP total	-.248	-.080	-.176	.450	.012	.073	.063	1	
Parental Care Total	.019	-.029	-.062	.070	.045	.354	.334	.275	1

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 9: Pearson correlations between IAT D scores and subscales of MORS, Parental Modernity (PM), parental stress and parental care total from the FAP

	1	2	3	4	5	6	7
1. D IAT Attachment/Leisure	1						
2. D reversed IAT my baby	.133	1					
3. RFQ certainty	.127	-.033	1				
4. RFQ uncertainty	.079	.132	-.330	1			
5. PRFQ Pre-Mentalizing Modes	.104	.080	-.292	.289	1		
6. PRFQ Certainty of Mental States	.093	.018	.413	-.268	-.074	1	
7. PRFQ Interest and Curiosity in Mental States	.207	-.175	.007	.010[**]	-.228[*]	-.067[**]	1

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 8: Pearson correlations between IAT D scores and measures of mentalizing: Reflective Function Questionnaire (RFQ) and Parental Reflective Function Questionnaire (PRFQ)

	1	2	3	4	5	6	7	8
1. D IAT Attachment/Leisure	1							
2. D IAT MyBaby	.133	1						
3. Activity level mean	.004	-.013	1					
4. Distress to Limitations	-.005	.019	.434	1				
5. Distress and latency to Approach Sudden or Novel stimuli	-.024	-.188	.243	.434	1			
6. Duration of orienting	.088	-.019	-.149	-.135	.042	1		
7. Smiling and laughing	.112	.030	.075	-.133**	-.210**	.273	1	
8. Soothability	-.092	.074	-.161	-.118	-.075	.056	.225	1

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Table 10: Pearson correlations between IAT D scores and IBQ subscales

Regression Analyses. *D* scores of each IAT were used as direct predictors of attachment security and attachment style in binary and multinomial logistic regression, respectively. Results were non-significant for each predictor and their interaction.

Mediation Analyses. *Indirect Effect of Implicit Attitudes on Infant Attachment Towards Parenting Through Mother's and Infant's Behaviour.* Behavioural data obtained by the CIB and the IBQ comprised 13 scales. These were reduced through principal component factor analysis. Factor loadings of <0.4 were suppressed. Both orthogonal and oblique rotations yielded a 3-factor solution explaining 60.66% of variance. The 3 factors grouped behaviour reported by observer, reported by mother,and a spurious factor whose components load more heavily on the other factors, with the exception of the "smiling and laughing" scale of the IBQ. This third factor will not be included in subsequent analyses (Table 11).

	Component		
	1.	2.	
	Observer-	Mother-	3.
	reported	reported	
CIB Dyad reciprocity	0.935		
CIB Parent sensitivity	0.886		
CIB Parent limit setting	0.865		
CIB Dyad negative states	-0.865		
CIB Child involvement	0.761		
CIB Child withdrawal	-0.697		
CIB Parent intrusiveness	-620		0.419
IBQ Distress to limitations		0.802	
IBQ Activity level		0.785	
IBQ Distress and latency to approach novel stimuli		0.545	-0.421
IBQ Soothability			
IBQ Duration of orienting			
IBQ Smiling and laughing			0.786

Table 11: Rotated component matrix for a factor analysis of behavioural measures of dyadic interaction. Extraction method: Principal component analysis. Rotation method: Varimax with Kaiser Normalisation

Mediation analyses were carried out, to determine the IATs' direct and indirect effect on attachment security through different measures of parenting behaviour.

The Attachment/Leisure IAT did not significantly predict attachment security as measured by the SSP. The only significant prediction was a direct effect of observed-rated parenting behaviour and attachment security when controlling for the effects of IAT and mother's report of child behaviour. The model and its significant regression coefficients are shown in Figure 3.
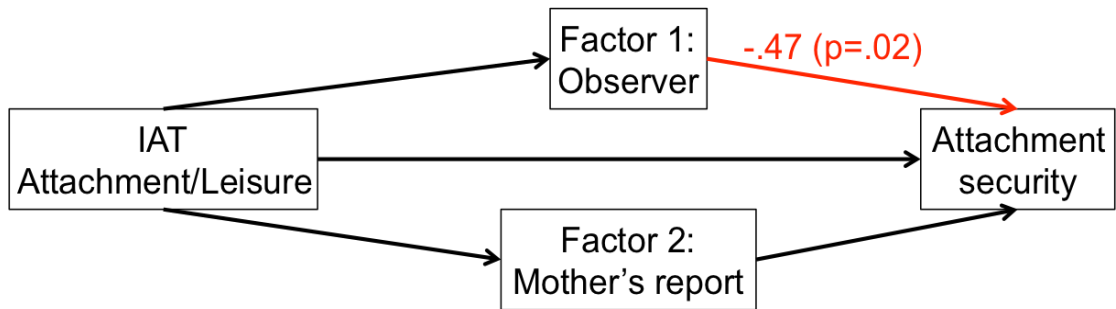
Figure 3: Mediation model for Attachment/Leisure IAT's effects on attachment security

Likewise, MyBaby IAT did not predict attachment security directly or indirectly through parent behaviour (Figure 4)

Figure 4: Mediation model for the effects of MyBaby IAT on attachment security

Indirect Effect of Implicit Attitudes towards Parenting on Infant Attachment Through Mothers' Conscious Attitudes Towards Parenting. Mediation analyses were carried out to establish direct and indirect effects of the IATs through different measures of mothers' explicit attitudes towards parenting:

- MORS Warmth-Coldness
- MORS Invasive-Withdrawn
- Progressive Beliefs about Raising Children
- Traditional Beliefs about Raising Children
- Ideas about Raising Children TOTAL
- Parental Stress

The Attachment/Leisure IAT showed a small but significant negative direct effect on Parental Stress when controlling for all other variables, but did not yield any significant direct or indirect effect on attachment security.

Measures of mother's warmth and coldness had a positive significant direct effect on attachment security (Figure 5).
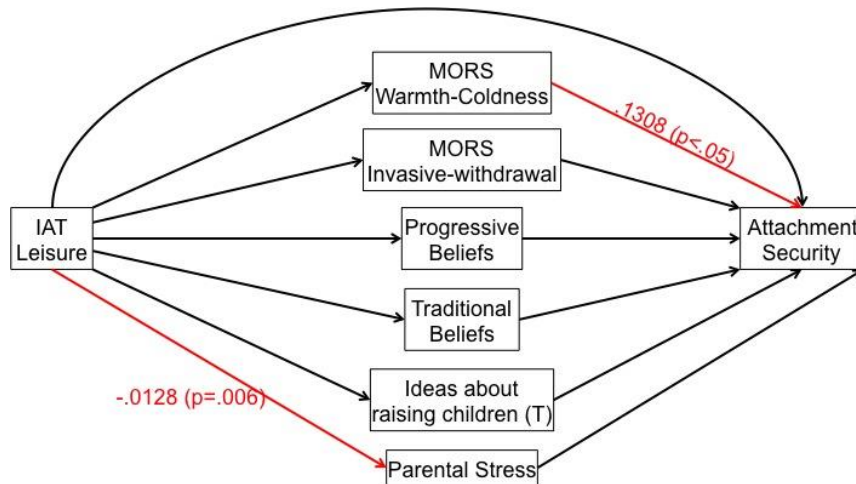
Figure 5: Mediation model for Attachment/Leisure IAT on attachment security

In the case of the MyBaby IAT, there were no significant associations found. The effect of mothers' warmth and coldness observed in the previous model only tends to significance (Figure 6).
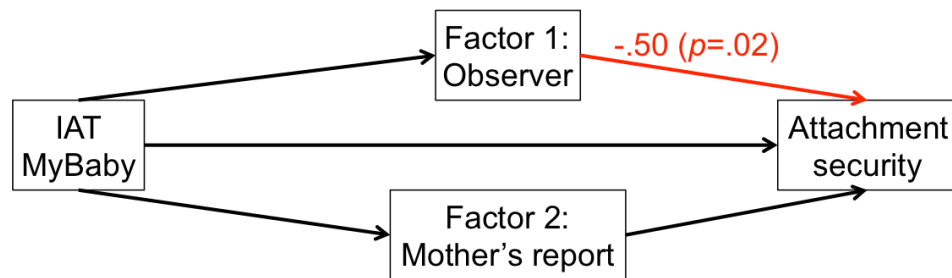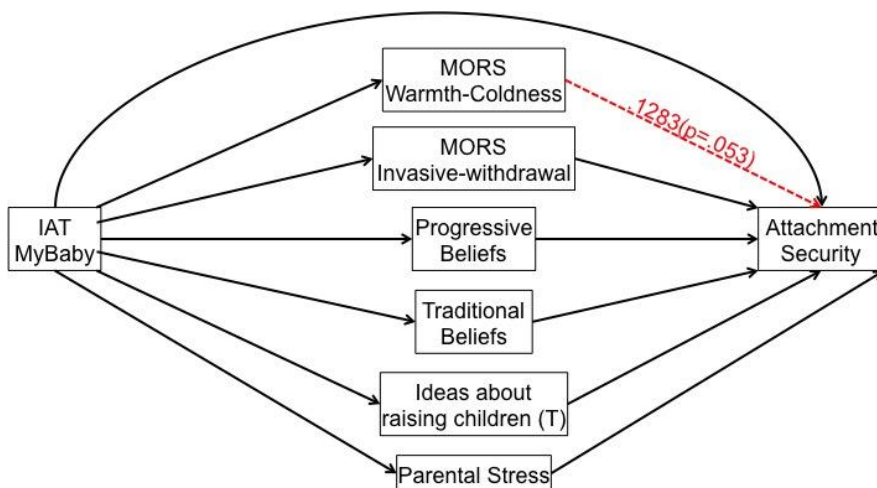

Figure 6: Mediation model for MyBaby IAT's effects on attachment security

Indirect Effect of Implicit Attitudes Towards Parenting on Attachment Security Mediated by Maternal Psychopathology. Factorial analysis of the BSI's subscales was carried out to reduce data for symptomatology. However, all subscales loaded onto one unique factor. Consequently, three parallel mediation analyses were carried out to test the effect of each IAT through the different BSI subscales grouped into internalising, externalising, and thought disorders. These three groups are broad, higher order structures of psychopathology both in children and adults (Caspi et al., 2014). The Attachment/Leisure IAT showed no significant effects in any of the three regressions. However, MyBaby IAT showed significant indirect effects on mothers' psychopathology for both anxiety (grouped as an internalising disorder) and psychoticism (thought disorder), but no indirect effects on attachment security (Figure 7 and Figure 8).

The rest of the BSI subscales showed no interactions in this model. There was no mediation of total scores of psychopathology in the regression of attachment security on MyBaby IAT scores.
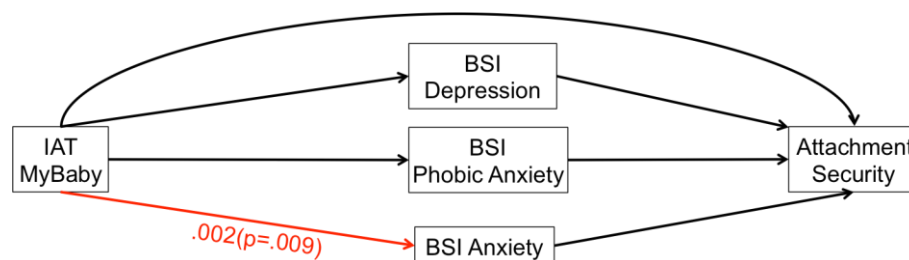
Figure 7: Mediation model of the effect of MyBaby IAT on attachment security mediated by internalising symptomatology
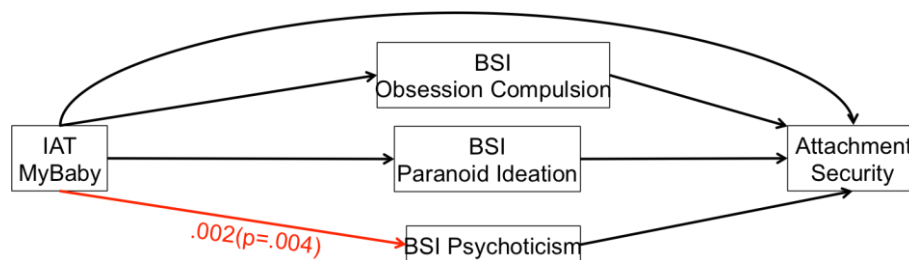
Figure 8: Effect of MyBaby IAT on attachment security mediated by thought psychopathology

Effects of Implicit Attitudes Towards Parenting on Attachment Security Mediated by Reflective Function. Regarding measures of reflective function, a similar model was devised to assess if the IATs can predict attachment security when mediated by reflective function. However, there were no statistically significant interactions of either IAT on attachment, directly or indirectly. Moreover, there were no direct effects of any of the variables. The model constructed to assess the same relationships, but this time using the PRFQ as a measure of reflective function, showed a significant direct effect of Interest and Curiosity in Mental States on Attachment Security (β = -1.02, p = 0.04). In the case of IAT MyBaby, the model yielded only a direct effect of Pre-Mentalizing Modes on Attachment, when controlling for IAT and other PRFQ subscale effects (β = 0.98, p = 0.05).

Discussion

In this study, in contrast to our expectations, IAT measures were unrelated to most constructs. Although we expected to find meaningful correlations between the IAT and questionnaire scores, their absence could be explained by the fact that both types of measures are believed to tap related, but different aspects of attachment representations and attitudes towards parenting (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005a; Maier et al., 2004). Studies assessing the relationship of implicit attitudes, as measured with the IAT and attachment seem to show interesting relationships. On the other hand, positive results combined with a scarcity of published articles might be an indicator of publication bias.

Both *D* scores were unable to differentiate between secure and insecure infants, or infants between attachment styles in the 3-way and 4-way classifications. This lack of significant results might be an indication of the small size of the sample when split into attachment style groups (Cohen, 1992).

The scarcity of significant results yielded by the data must be interpreted with care. Firstly, they run against theoretical expectations; secondly, the associations found are weak and unrelated to measures of attachment security and attachment style. These effects, however weak, associated more positive implicit attitudes towards parenting and attachment issues with measures of anxiety. The MyBaby IAT, the one measuring the preference of the mother's own baby in comparison with another baby, was associated with both anxiety and psychoticism, which were in turn strongly correlated

with each other, given that it was a healthy sample where symptomatic scores tended to zero.

Similarly highlighting anxiety, the relationship between parental stress and implicit positive attitudes towards attachment pictures was also significant. These relations could be explained by the mother's differential sensitivity to relational cues, typically observed in anxious people (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & Van Ijzendoorn, 2007; Fox, Mathews, Calder, & Yiend, 2007). This same anxiety could explain the relationship between the MyBaby IAT scores and the distress and latency subscale of the IBQ, which is a measure completed by mothers according to their observations of their infant's behaviour. As it has been seen in other studies (see Shaver & Mikulincer, 2002a), anxious mothers could be rating their babies as more distressed than they really are by "projecting" their own anxiety. However, it is impossible to know how these factors influence the transmission of attachment, given that these measures had no relationship to babies' attachment and that mothers' attachment was not measured.

These unexpected results have a two-fold consequence for the rest of this doctoral thesis. On the one hand they fail to yield evidence about a relationship between implicit attitudes towards parenting and attachment, and offspring's attachment status. On the other hand, they highlight the association between implicit attitudes and measures of psychopathology. This relationship is very interesting to explore and it will be more deeply addressed from Chapter 4 onwards.

Limitations

An important limitation of this study might be related to the observation that some mothers found it awkward to complete the IATs. Hence, further work concerning its validity is needed.

The sample had an increased proportion of securely attached infants, which impeded reliable group comparisons.

Among the other limitations of this study, the lack of a measure assessing mothers' attachment style seems to be the most important. For instance, mothers who showed increased anxiety may have had a preoccupied attachment style, which would fit the hypotheses of a hyperactivation of the attachment system of these mothers and the concomitant oversensitivity to parenting cues (Adam et al., 2004). It is well demonstrated that more anxious mothers attribute less importance to the development of

their children's independence, perhaps explaining the relationship of the IAT with the scale of parental intrusiveness (Mikulincer & Shaver, 2012).

Parent attachment style has also been closely related to the quality of parenting (Adam et al., 2004; Green et al., 2007). Attachment-related differences in providing care to children are already evident in young adults' pre-parenting expectations. People who score high on either attachment anxiety or avoidance are less positive than their more secure peers when it comes to judging their ability to relate to children and imagining relationships with their own future children. Additionally, while more avoidant people are less interested in having children and anticipate less satisfaction from caring for them, more attachment-anxious individuals hold unrealistic, perfectionist expectations, which may reflect wishes to be loved and to overcome self-doubt about caregiving skills.

Similar findings have been reported about people who are parents already (Mayseless & Scher, 2000; Mikulincer & Shaver, 2012). It is important to remember here that, notwithstanding the close relationship between the attachment system and the caregiving system, these are two independent behavioural conglomerates, and that the variations in one do not explain the variations of the other entirely. Life circumstances can obstruct the provision of care (Green et al., 2007).

The selection of GOOD and BAD words to be used as the IAT's stimuli were rated by experts as pertaining to one of four domains, namely the four attachment styles: secure, avoidant, preoccupied and disorganised. However, they were used as if pertaining to only two groups i.e. GOOD and BAD. This could have affected the IAT's results, given that the appearance of a word related to anxiety or avoidance could have shown a stronger association to attachment in mothers presenting these insecure styles. As mentioned in the introduction, current research and theory argues that differences in attachment should be organised around either the pole anxiety/avoidance and the pole self/other (Dewitte et al., 2008; Veletanlic, 2007; Zayas & Shoda, 2004; Zayas & Shoda, 2005). The construction of the IATs used in this study does not target those dimensions because they are new adaptations of the measure.

The IAT that yielded less results showed pictures of babies being comforted by an adult (ATTACHMENT) or pictures of an adult engaged in an activity by him/herself (LEISURE). Some of these activities, as seen in Figure 1, are actually quite pleasurable, which implies their possible uselessness as contrasts to ATTACHMENT and a questionable BAD status.

A last consideration is that automatic attachment evaluations were measured in a relatively neutral and stress-free context. This is potentially problematic because attachment theory highlights the role of internal working models in regulating proximity and felt security when confronted with distress. From this idea it can be deduced that individual differences in implicit attitudes towards attachment could be made more salient if the images contained infants in distress.

Chapter 3: Self-Esteem IAT for iPad: Discriminant Validity and Reliability Analyses in a Sample of Healthy Adolescents

Introduction

The present chapter introduces a new version of the Self-Esteem Implicit Associations Test (SE-IAT), an adaptation of the SE-IAT used by previous studies in the literature (Steinberg, 2006; Steinberg, Karpinski, & Alloy, 2007) to a portable tablet format, namely an iPad. The main objective of the present study is to establish psychometric characteristics of this adaptation and to ascertain its robustness for variations in gender and age, in preparation for the next series of studies in this thesis, which utilise this adaptation of the SE-IAT to predict measures of depression.

The purpose of this validation study is to introduce the reader to the specific processes of validation for a time-reaction test that relies on discrepancy scores in a normative sample. It addresses particularly the topics of age and gender, given the variability of the sample in those demographic variables. These validation methods will then be used for the rest of the thesis, but their detailed introduction and explanation is the main goal of this study. Given that it is a normative sample, implicit self-esteem scores obtained with the IAT will not be contrasted against measures of psychopathology, which is the topic of Chapters 4 and onwards.

In order to introduce this study, this chapter starts with a brief revision of the status of the concept of self-esteem in contemporary psychological research, particularly in the field of development and gender differences. It then gives an overview of the psychometric properties of the IAT in the literature.

The Study of Self-Esteem

The concept of self-esteem is one of the few constructs of psychological science that has a place in most formulations of psychopathology, psychotherapy and personality (DeHart, Peña, & Tennen, 2013). It constitutes an essential concept in the study of highly prevalent disorders such as depression, mania and borderline personality disorder. Not only its trait level, but its fluctuations are also central to the understanding of narcissistic personality disorder and eating disorders (Geller et al., 1998; Gual et al., 2002; Kohut & Wolf, 1978; Salman Akhtar & Thomson Jr, 1982; Watson, Little, Sawrie, & Biderman, 1992; Williams et al., 1993).

As a concept within psychological science, it has its origin in the work of William James at the end of the 19th Century. He differentiated three domains of the self, namely the material, the social and the spiritual. Self-esteem belong to the social self, that is to say, those characteristics of oneself that others can perceive (James, 1984). Self-esteem is an attitude, namely the subjectively perceived value of the self. Therefore it has been studied with the same methods used in general attitudinal research.[5]

Most of the research on self-esteem has focused on people's explicit self-evaluations. Amongst the frequently used measures are those identified by Blaskovich and Tomaka (1991): The Coopersmith's Self-Esteem Inventory, the Tennessee Self-Concept Scale, Janis and Field's Feelings of Inadequacy Scale, and the Texas Social Behavior Inventory, to which Bosson (2006) adds the Self-Liking and Self-Competence Scale and the State Self-Esteem Scale. The most widespread measure to assess self-esteem is the Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965). It is a 10-item self-report measure, which asks people to reflect upon their self-worth and feelings about themselves. It was originally developed to be used with adolescents, but since its creation it has been utilised in many contexts and populations. Example items are: "On the whole, I am satisfied with myself" and "At times I think I am no good at all". This measure yields one score of global self-esteem. However, further investigations have shown a two-factor structure of the global self-esteem construct (Tafarodi & Milne, 2002). One reflects one's own moral significance (*I am good* or *I am bad)*, with items such as, "I take a positive attitude toward myself"; the other is a reflection of the experience of one's own power or efficacy (*I do things well* or *I do things badly)*, with items like "I am able to do things as well as most other people" (Tafarodi, 2006, p. 111).

Asking people directly how much they value themselves has proven very fruitful in research (Karpinski & Steinberg, 2006). The aforementioned measures of self-reported self-esteem are able to predict various behaviour and outcomes such as psychological well-being, relationship and life satisfaction, academic achievement and even physical health (Baumeister et al., 2003; Bednar & Peterson, 1995; DuBois, Felner, Brand, & George, 1999; Dubois & Tevendale, 1999). Furthermore, these measures are inexpensive and need little training in administration and scoring, which

---

[5] It is noteworthy that in spite of the various psychological outcomes related to self-esteem, it is its subjective element that differentiates it from actual success and performance in life. Self-esteem is not related to people's objective talents or achievements, as shown in a review of more than 20,000 participants (Baumeister, Campbell, Krueger, & Vohs, 2003).

implies that they can be used in multiple contexts and with different populations. Most of these self-report measures demonstrate strong psychometric properties and show robustness to demographic factors like gender, age and marital status (Bosson, 2006; Bosson, Swann, & Pennebaker, 2000; Fleming & Courtney, 1984). Most importantly, the high face validity of self-report measures of self-esteem facilitates the interpretation of results by researchers (Koestner & Mageau, 2006).

However, there is a major limitation to the study of explicit and declarable self-esteem. As with other attitudes, when people are asked directly, they have great control over their responses. Self-presentation biases might be influencing the way that people answer those direct probes, or it is possible that in spite of an effort to be honest, there are aspects of self-esteem that are not available to introspection (Greenwald & Banaji, 1995). It has also been observed that high explicit self-esteem serves a defensive function: if confronted with ideas about one's own death, one is more likely to show increased self-esteem (Harmon-Jones et al., 1997; Pyszczynski, Greenberg, Solomon, Arndt, & Schimel, 2004). In sum, it is logical to doubt that self-reported self-esteem actually reflects the full and true feelings of the respondents. This is particularly problematic for clinical research because psychopathology has long been conceptualised as operating outside conscious awareness (Beck & Alford, 1967; Breuer & Freud, 1895).

Implicit study of self-esteem. Self-esteem is one of the concepts that has benefited from the emergence and growing popularity of implicit measures of attitudes reviewed in Chapter 1. Several measures have been developed to attempt to measure aspects of self-evaluation not readily conscious and free from self-presentational biases. That is to say, to measure implicit self-esteem: an automatic, overlearned, and non-conscious evaluation of the self that guides spontaneous reactions to self-relevant stimuli (Greenwald & Banaji, 1995). These spontaneous reactions can be operationalised as differences in reaction times or preferences for ideas and concepts apparently unrelated to the self, but which may mask a hidden relationship with the self. This is because, on average, people tend to evaluate stimuli associated with the self more positively than those with no association, while it is still possible to find individual differences within that tendency (Greenwald & Banaji, 1995). Thus, when people are not necessarily aware of this evaluative tendency, it can be conceived as reflecting implicit attitudes towards the self. For example, people show a preference for the initials of their names over other letters of the alphabet, which can be understood as a variant of the classic halo effect (Hoorens & Nuttin, 1993; Thorndike, 1920).

Regarding reaction times as an operationalisation of implicit self-esteem, research shows that the mere encounter with an object will automatically and effortlessly elicit an evaluation, a phenomenon known as the "automatic attitude activation effect" (Bargh, Chaiken, Govender, & Pratto, 1992, p. 893; Fazio et al., 1989). It is then logical to think that the encounter of a person with him/herself would also elicit such an automatic evaluation. Once activated, this evaluation facilitates the processing of similar stimuli while impeding the processing of dissimilar stimuli. This affective evaluation can then be seen in how people show faster responses when faced with stimuli that are consistent, in comparison with those that are inconsistent. If someone whose self-evaluation is positive pairs positive words faster with the concept of the self, than they would when the self is associated with negative-valence words, or if a person associates holidays with fun, and they are presented with the word "holidays", they will subsequently be faster at recognising the word "happy" than the word "sad" (Bosson et al., 2000). The extent to which this process is facilitated when related to the self is taken as a measure of implicit self-esteem.

Among the various methods used to measure implicit self-esteem, one of the most popular, together with the IAT, is the Name Letter Task (Nuttin, 1985). Participants are required to rate their liking for each letter of the alphabet. As explained above, high implicit self-esteem is indexed by the extent to which a person prefers his or her initials to other letters of the alphabet. The Supraliminal Attitude-Prime Task (Hetts et al., 1999) is a time-reaction measure based on priming methodologies. Participants in front of a screen see a series of words related to oneself (e.g. me), and after that, a positive or negative word is presented. People must then classify that second word as good or bad, as quickly as possible. They repeat this procedure 20 times. An explicit self-esteem index is obtained by comparing the time reactions of the pairing of "me" with good words and the pairing of "me" with bad words. A similar measure presents all words subliminally during 96 trials (the Subliminal Attitude-Prime Task; Spalding & Hardin, 2000). A modification of the Stroop Task has been developed, in which participants are presented with positive, negative or neutral self-related stimuli printed in different colours on a screen. Participants must then indicate the colour of the phrase presented as fast as possible. A measure of implicit self-esteem is then obtained when time-reactions show facilitated responses to positive self-relevant phrases (Bosson et al., 2000). Another popular indirect measure of self-esteem is the Ambiguous Statements Task, which requires the participant to imagine an acquaintance saying 13 different ambiguous sentences like "What did you say?" (Tafarodi, 1998, p.1190), and are

required to rate the extent to which these phrases reflect positive or negative feelings towards them.

These are the most popular measures in the literature, other measures are the use of word fragments and sentence completions (some of these under cognitive load), the extrinsic affective Simon Test, the Go/No-Go association tests, and measurements of signature size (Koole et al., 2007). When these various measures of implicit self-esteem are compared, results are mixed. Bosson et al (2000) found that correlations between these measures are at best low. However, several meaningful findings have been replicated across different measures of implicit self-esteem. For example, exposing a subject to repeated associations of the self with positive characteristics causes a temporal increase for scores of implicit self-esteem in both the IAT and the Name Letter Task (Baccus, Baldwin, & Packer, 2004; Dijksterhuis, 2004). This implies that these measures tap into similar psychological processes, or various aspects of the same process, despite their low inter-correlations (Koole et al., 2007).

Explicit and implicit self-esteem are, at best, only weakly correlated to each other. Furthermore, these two aspects of self-esteem have discriminant predictive validity: implicit self-esteem is a better predictor than explicit self-esteem of spontaneous and affectively driven behaviour, such as a depressive mood in response to threatening feedback (Greenwald & Farnham, 2000), and can even influence major life decisions including choice of spouse, career and residence (Greenwald & Banaji, 1995; Jones, Pelham, Mirenberg, & Hetts, 2002; Pelham, Mirenberg, & Jones, 2002).

Development of self-esteem. The study of explicit self-esteem development during adolescence and across the life span in general yields conflicting results. Studies report that while self-esteem is relatively high during childhood, in adolescence it is possible to observe a drop in the average of self-esteem scores, especially for females (Robins & Trzesniewski, 2005; Steer, Ball, & Ranieri, 1999). A cross-sectional study of people aged 9-90 showed the same pattern (Robins, Trzesniewski, Tracy, Gosling, & Potter, 2002). Concurrently, other studies show that, during adolescence, self-esteem increases. This increase becomes slower during young adulthood (Erol & Orth, 2011). Longitudinal data on 1,824 individuals shows a somewhat similar pattern of self-esteem increase during adolescence and adulthood, reaching a peak at the age of 50 and decreasing with old age (Orth, Robins, & Widaman, 2012).

Notwithstanding these dissimilar results, there is enough evidence to support the notion of longitudinal rank-order stability of self-esteem. That is to say, individuals who

have relatively high self-esteem at one point in time, tend to show high self-esteem years later (Orth & Robins, 2013).

As with its explicit counterpart, implicit self-esteem also originates and is shaped through interactions with others. Theoretically speaking, both conscious and unconscious ideas of the self in relation to others are formed early in life, based on the relationship between a child and his or her caregivers (Bowlby, 1969, 1973). It is not surprising then to find that a secure attachment style measured in infancy is capable of predicting higher explicit self-esteem at pre-school age.(Dykas & Cassidy, 2011; Sroufe, 1983). In this same line, more positive implicit self-esteem in adulthood (measured with the Name Letter Task) is related to self-reports of having a mother that was more nurturing and less overprotective during childhood. These results are based on subjects' recollections of their own childhood, and were maintained when their mothers independently reported their early interactions with their children. The association between parenting styles and implicit self-esteem occurred over and above explicit self-esteem (DeHart, Pelham, & Tennen, 2006; DeHart et al., 2013).

There are no further examples of implicit self-esteem development in the literature, but it is logical to conclude that if implicit self-esteem is an overlearned automatic valuation of the self over which there is limited conscious control, its level remains relatively stable during the life span. It has been argued in the literature that even though both explicit and implicit dimensions of self-esteem have social origins and develop based on interactions with significant others, they are established in distinct stages of life. Although literature regarding explicit self-esteem is contradictory regarding the direction of increments and decrements of self-esteem during the life span, there is consensus regarding the fact that it does vary during later stages of life, which is not the case for implicit self-esteem.

Gender differences in self-esteem. A study that shows an increase in self-esteem during adolescence, in a cross-sectional sample of 7,100 individuals aged 14-30 showed that the developmental trajectories of self-esteem are the same for males and females (Erol & Orth, 2011). However, a meta-analysis with more than 100,000 participants showed that females have significantly lower self-esteem during adolescence and early adulthood, compared to adolescent and early adult males (Kling, Hyde, Showers, & Buswell, 1999). In this same line, other studies found that, during adolescence, self-esteem declines in women, but increases in men (Block & Robins, 1993; Zimmerman, Copeland, Shope, & Dielman, 1997). Self-esteem for adolescent women is more

focused on appearance, with women who are dissatisfied with their bodily appearance showing lower self-esteem. On the other hand, for adolescent men, self-esteem is more related to athletic performance, but in a different way: males who are satisfied with their athletic performance show higher self-esteem, but there is no necessary relation between dissatisfaction and lower self-esteem. That is to say, adolescent males' satisfaction with their bodies increments self-esteem levels, but dissatisfaction does not necessarily entail a low level of self-esteem (Bolognini, Plancherel, Bettschart, & Halfon, 1996; Furnham, Badmin, & Sneade, 2002). There is a trend in the literature that finds small differences in self-esteem according to gender, with males presenting higher levels of explicit self-esteem than females, the largest difference being found in late adolescence (Kling et al., 1999).

In the case of implicit self-esteem, 6 related studies (Pelham et al., 2005) set out to explore the relationship between implicit and explicit self-esteem in men and women. The authors found that gender moderates the relationship between these two levels of self-esteem, based on the idea that women are more strongly socialised to trust their feelings and intuitions. All the studies showed that explicit and implicit self-esteem relate more to each other in women than in men, using the Name Letter Task and the Implicit Self-Evaluation Scale. The results were replicated for both of the implicit self-esteem measures and in three different cultural contexts, namely a Dutch sample, a North American (California) sample and a Singaporean sample. But such effect is not necessarily dependent on gender only. Both males and females who trust their intuition show a larger correspondence between implicit and explicit attitudes than people who do not trust their intuition (Hofmann, Gschwendner, Nosek, & Schmitt, 2005b).

Most studies addressing gender effects on implicit self-esteem make use of the IAT, therefore these will be reviewed below, when I summarise the findings regarding demographic effect on the use of this measure.

Demographic Differences in the Use of the IAT

The Implicit Associations Test, as we have already seen, has been adapted to measure various types of attitudes. Gender differences emerge scarcely and are normally related to the attitudinal object being measured. For example, a study showed no gender differences in implicit self-esteem scores as measured with the SE-IAT, but it did show a marginal gender difference in the way these scores were predictive, together with explicit self-esteem, of social anxiety in women only. The association between low

implicit self-esteem and social anxiety was most evident for women with relatively low explicit self-esteem (De Jong, Sportel, De Hullu, & Nauta, 2012).

A similar phenomenon is observable in the administration of an IAT measuring implicit neuroticism and attachment anxiety. Attachment anxiety was correlated with neuroticism as measured by the IAT in women but not in men (Donges, Jachmann, Kersting, Egloff, & Suslow, 2015).

Within these small gender differences, when the IAT measures the attitudes towards gender, women tend to show more preference for their in-group (i.e. other women), than men towards other men (Aidman & Carroll, 2003; Rudman & Goodwin, 2004). Men demonstrate stronger implicit preferences in IATs measuring attitudes to social groups: regarding thin versus fat people, even overweight men showed a stronger preference towards thin people, while that effect was not present for women. This can imply that while women show stronger preferences towards their in-group, men show stronger preferences toward groups with higher social status. (Nosek et al., 2002). With regards to political attitudes, when gender is added alongside ethnicity and explicit political orientation using a model predicting implicit political attitudes, the predictive power of gender is rendered non-significant. These and the following results were obtained from a massive sample of more than 2.5 million IATs spanning 17 different attitudinal objects (Nosek et al., 2007b).

Regarding age, little psychological research has addressed variation in social group attitudes in the IAT according to the age of the respondent. Samples of all ages show a positive implicit preference for younger rather than older faces, similar implicit attitudes are found toward Judaism, weight, and race. When age differences do appear, they are consistent but small: older individuals show greater negativity towards dark-skinned, Arab-Muslim people, and disabled people than young adults. A curved relationship between the respondent age and implicit attitudes appears in another two domains: the youngest and oldest extremes of IAT respondents show implicit negativity towards black people and homosexual people, while middle age respondents do not show such negativity. Older people tend to show more automatic social stereotypes: they more readily associate women with family and men with career, men with science and women with humanities. These variations were also observable in explicit measures (Nosek et al., 2007b).

The IAT's Psychometric Properties

As with any other instrument used to measure individual differences in psychology, establishing good psychometric properties is essential for the IAT: it is required that the measure shows stability of implicit self-representations over time, and that these contribute to the prediction of behaviour in a significant and independent manner, without covarying confounding variables (like demographic characteristics of the sample). Regrettably most of the investigations using the IAT do not report on the psychometric properties of the instrument (Gawronski, 2009).

Among the various measures of implicit self-esteem used by researchers, only two achieve acceptable test-retest reliability indexes (for an average time span of 31.23 days; minimum 22 days, maximum 38 days), namely the SE-IAT (0.69) and the Name Letter Task (0.63). This is very high in comparison with the Supraliminal Attitude-Prime Task (0.08), Subliminal Attitude-Prime Task (0.28), the Stroop Task (-0.05), and other implicit measures of self-esteem (Bosson et al., 2000). Other examples in the literature show a median of 0.56 for test-retest reliabilities in different versions of the IAT, including time spans of minutes to one year (Nosek et al., 2007a). Despite these indexes being superior to all other implicit measures to which it the IAT has been compared, when contrasted with the most popular measure of explicit self-esteem, the RSES (0.80), these estimates are low (Bosson et al., 2000).

As I have already stated, most of the implicit self-esteem measures used in the literature correlate only weakly with other measures of implicit self-esteem, which might be interpreted as a good indicator of discriminant validity (Olson & Fazio, 2003) but might also be reflecting the unacceptable psychometric properties of other implicit measures to which the IAT is compared: When controlling for the lack of other measures' reliability, correlations with the IAT are increased (Cunningham, Preacher, & Banaji, 2001). Convergent validity of the IAT with a similar procedure that uses joystick movements towards or away from the participant instead of pressing keys is much improved (0.50 with the Implicit Association Procedure; see Schnabel, Banse, & Asendorpf, 2006). These correlations, which are at best medium, indicate an upper bound to construct validity and suggest a substantial method variance in implicit procedures in general.

Correlations between the IATs and different explicit measures, including explicit self-esteem are weak and positive, and most of the time only marginally significant (Koestner & Mageau, 2006; Teige, Schnabel, Banse, & Asendorpf, 2004). However,

convergent validity of implicit self-esteem measures has been only demonstrated by the SE-IAT, which is able to predict, as other explicit measures do, the ratings of independent evaluators who were presented with essays of SE-IAT responders about their self-views (Bosson et al., 2000). The predictive value of the IAT is, in many cases, in the same direction as (but independent of) explicit measures, in spite of correlations between the measures being small (Back et al., 2009). In other studies, the predictive validity of self-esteem IATs yields indicators of interactive validity effects, namely implicit and explicit self-esteem interact to predict outcomes. For example, people with discrepancies between implicit and explicit self-esteem exhibit more defensive behaviours like rejection of negative feedback or exaggerated social consensus estimates (Jordan, Spencer, Zanna, Hoshino-Browne, & Correll, 2003; McGregor, Nail, Marigold, & Kang, 2005; Rudolph, Schröder-Abe, Schutz, Gregg, & Sedikides, 2008; Schröder- Abé, Rudolph, Wiesner, & Schütz, 2007b). These results evidence both convergent and discriminant validity: the IAT and its explicit counterparts address related but distinct constructs (Nosek & Smyth, 2007; Nosek et al., 2007b). It is important to note that the relationship between implicit and explicit measures varies according to the attitudinal object measured. Other variables that moderate the relationship between implicit and explicit measures can be gender (see above Pelham et al., 2005), the relationship and similarity between explicit and implicit items and words (or other stimuli) used, the spontaneity in the generation of self-reports, method-specific variance of both implicit and explicit measures (e.g. fakability), sampling biases, and social desirability (Hofmann et al., 2005b). The effect of social desirability has been found to be negligible in the case of self-concept (Egloff & Schmukle, 2003).

Different adaptations of the IAT usually reach estimates of internal reliability between 0.70 and 0.90, (Banse, Seise, & Zerbes, 2001; Bosson et al., 2000; Nosek et al., 2007a; Schmukle & Egloff, 2004), which represent good to excellent levels (Kline, 2013). Besides being psychometrically satisfactory, these indexes of internal consistency are the highest found among implicit measures, in comparison, for example, with supraliminal and subliminal priming measures (Bosson et al., 2000; Kawakami & Dovidio, 2001), and the Go/No-Go Association Task (Nosek & Banaji, 2001). Given the large amount of probes that a typical IAT has (120 in the SE-IAT used in this thesis, see below for the description of our measure), and the fact that it is not composed of items, but of difference scores, several methods for calculating internal consistency have been used throughout the literature. Some authors compute difference scores for each single trial of the incompatible condition minus its correspondent trial of the

compatible condition (according to presentation order), and use them as single "items" to calculate Cronbach's α (Cunningham et al., 2001). However, this blockwise comparison might be affected by factors extrinsic to the measure (e.g. participant's fatigue will imply an upward or downward trend in difference scores as later trials will show higher response latencies). Other authors employ difference scores for (sometimes randomly) selected groups of 5, 10 or more trials to calculate α (Egloff, Schwerdtfeger, & Schmukle, 2005; Gawronski, 2002). According to Schmukle and Egloff (2004). A more useful way to calculate split-half reliabilities for the IAT is to separate difference scores, according to presentation order, in odd vs. even trials, therefore controlling for response changes during completion of the measure. The standard procedure (Schnabel et al., 2008) is to calculate split-half reliabilities over the difference scores of blocks 5.1 - 3.1 and 5.2 - 3.2 (see Table 14). In this chapter, all these calculation methods will be used.

In spite of the fact that the different IAT measures are less prone to faking responses than their explicit counterparts (that being one of the main reason for the use of the IAT; Banse et al., 2001; Steffens, 2004), they are still marginally susceptible to faking. It has been demonstrated that when instructed on how to fake, this susceptibility increases (Kim, 2003). But faking only compromises the validity and usefulness of the IAT to measure group differences when different individuals fake to a differing extent, which usually does not occur in experimental settings (a phenomenon named "differential faking"; Schnabel et al., 2006, p. 70).

In sum, the IAT in all versions appears to have the best psychometric properties among most implicit measures. In particular the SE-IAT appears uncontaminated by demographic characteristics of the sample.

Hypotheses

Taking into account this brief revision of the literature, it is expected that implicit self-esteem in normally developing adolescents and young adults will remain stable across gender and age. This will show the discriminant validity of this version of the SE-IAT and its robustness to these demographic characteristics. It is expected that the SE-IAT will show good indexes of internal consistency, as other versions of the measure have proven.

This is the first study in this doctoral thesis which utilises the adaptation of the Self-Esteem Implicit Association Test to a tablet computer format, therefore good

psychometric results will determine whether the test functions as expected, and if it is sensible to keep using this novel version in the rest of the studies of this thesis.

Methods

Design

The present is a cross-sectional correlational study. In order to test the hypotheses, the SE-IAT was administered to a sample of healthy female and male adolescents of various ages. IAT D scores are the dependent variable, while demographic characteristics are independent variables. This design does not contemplate the study of relationships between the SE-IAT and other measures collected by U-CHANGE (see below).

Procedure

The present study is framed within a larger longitudinal and accelerated longitudinal study, *Understanding and Characterising Health Adolescent-to-Adult Neurodevelopmental Growth Effects* (U-CHANGE). This larger study, to be completed in March 2017, aims at following a large sample of healthy adolescents (3,000 approx.) aged 14-24, in order to establish the neurocognitive unfolding of normal development. It forms part of the strategic planning of the Neurosciences for Psychiatry Network (NSPN) in the United Kingdom, and it will build a large normative database that will support future neuroscience-driven studies of adolescents and young adult patients with past or current mental disorders. The neuroscientific data obtained by this study will be related to genetic phenotypes and to endocrine measures. The larger study functions at

ME stimuli words	NOT ME stimuli words
I	they
me	them
my	their
mine	it
self	other

Table 12: "me" and "not me" stimuli words for the SE-IAT

two sites: the Anna Freud Centre in London and the University of Cambridge, Cambridgeshire.

The sub-sample for the present study comprises of participants from London, measured at the Anna Freud Centre.

After completing the SE-IAT, participants go on to complete a group of 7 age-appropriate cognitive computer tasks, a clinical interview, a test of language proficiency, measures of height and weight, and provide saliva swabs for genetic testing. The duration of one IAT testing session is approximately 5-7 minutes.

Participants

Participants were recruited through advertisement in schools, universities, colleges and young people's social clubs. The recruitment was also advertised in a high-circulation newspaper (Metro). Participation was strictly voluntary, but subjects were paid for their time and travel expenses.

Measures

Self-Esteem Implicit Associations Test (SE-IAT)

In this study, we have adapted the SE-IAT from its traditional desktop computer version, where the participant must sort the stimuli pressing the keyboard, to an iPad version, where participants must tap on either side of the screen to categorise the words presented. This version of the SE-IAT will be used in all the following studies presented in this thesis. Stimuli words for both the "me vs. not me" and the "positive vs. negative" categories are shown in Table 12 and Table 13 (Steinberg et al., 2007).

All participants completed the SE-IAT on an iPad 2 by Apple, which was connected to a centralised research database at UCL, the Patient Owned Database (POD).[6] This database hosts the measures in a digital form and securely stores and organises scores from various on-going research projects, anonymously and in real time. The SE-IAT was adapted to POD using JavaScript. The script was tied into the computer clock to ensure accurate timing. The word stimuli were stored in an array and drawn out to the front-end using a randomiser. The Java applet was designed to respond to a screen being touched instead of keyboard depression. Finally, the anonymised data was stored in POD and was retrievable only by authorised researchers, who received authorisation according to participant consent and study design. The data was retrieved in a Comma Separated Value (*.csv) file, readable by spreadsheet programmes like

POSITIVE words	NEGATIVE words
smart	stupid
bright	ugly
success	failure
splendid	awful
valued	useless
noble	vile
strong	weak
proud	ashamed
loved	hated
honest	guilty
competent	awkward
worthy	rotten
nice	despised

Table 13: "positive" vs. "negative" stimuli words for the SE-IAT

[6] http://www.ucl.ac.uk/pod

Step 1 Practice block (20 trials)	NOT ME		ME
	○	self	◉
	◉	other	○

Step 2 Practice block (20 trials)	NEGATIVE		POSITIVE
	○	joy	◉
	◉	vomit	○

Step 3 3.1 Practice block (20 trials) 3.2 Critical block (40 trials)	NEGATIVE OR NOT ME		POSITIVE OR ME
	○	self	◉
	○	joy	◉
	◉	other	○
	◉	vomit	○

Step 4 Practice block (20 trials)	POSITIVE		NEGATIVE
	◉	joy	○
	○	vomit	◉

Step 5 5.1 Practice block (20 trials) 5.2 Critical block (40 trials)	NEGATIVE OR ME		POSTIVE OR NOT ME
	◉	self	○
	○	joy	◉
	○	other	◉
	◉	vomit	○

Table 14: SE-IAT sequence. Full circles represent correct responses for the example trial. Adapted from (Greenwald & Farnham, 2000).

Microsoft Excel.

Participants completed the IAT before the rest of the self-report questionnaires. After seeing a screen with instructions explaining that they had to sort out words to different categories as fast as possible, participants taped a "next" button and started the test. This was formed of seven blocks showed in Table 14 (Greenwald & Farnham, 2000).

During steps 1, 2 and 4, participants had to rapidly categorise words appearing on the centre of the screen in 20 trials per step. These words were randomly selected from the "me" and "not me" lists, and from the "positive" and "negative" lists (Table 12 and

Table 13) they had to tap on either side of the screen, to match it with the category "ME" or "NOT ME" and "POSITIVE" or "NEGATIVE" appearing at the top of each side of the screen (Figure 9). During step one, participants categorised *target concepts* (me vs. not me); in steps 3 and 5, participants had to categorise *attributes* (positive vs. negative).
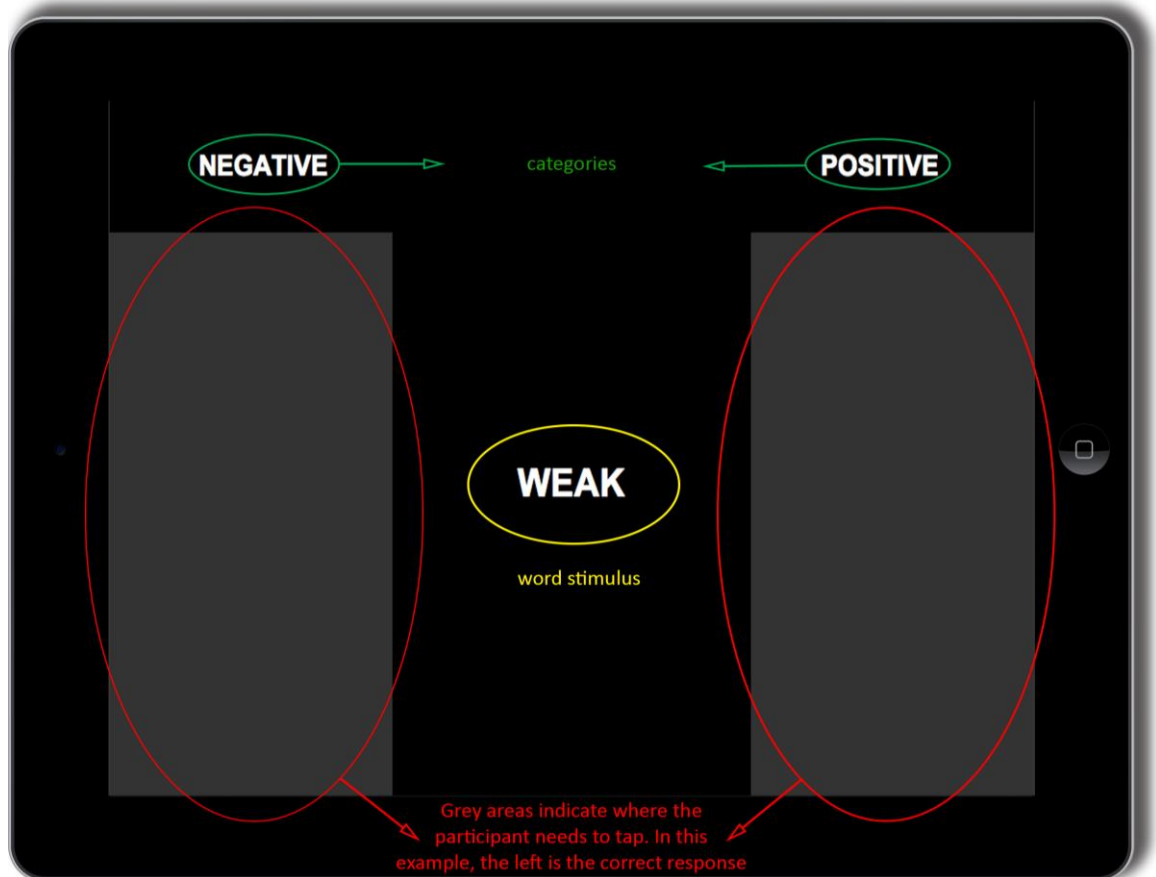
Figure 9: Trial layout taken from Step 2. Green, red and yellow circles, arrows and text have been added to this image for explanatory purposes.

During steps 3 and 5, participants had to rapidly categorise words presented in the middle of the screen, but now the categories are combined, each including one target concept and one attribute (Figure 10). Step 5 is the inverted form of Step 3, where concept and attribute are combined in different ways.

In each step, if the participant taps the wrong side of the screen, a red "X" appears in the middle of the screen and disappears only when the participant taps the correct side. This second attempt is not timed, but the software is capable of recording whether the first attempt was correct or not. Between each word and the next, there is an interstitial time of 400 ms.
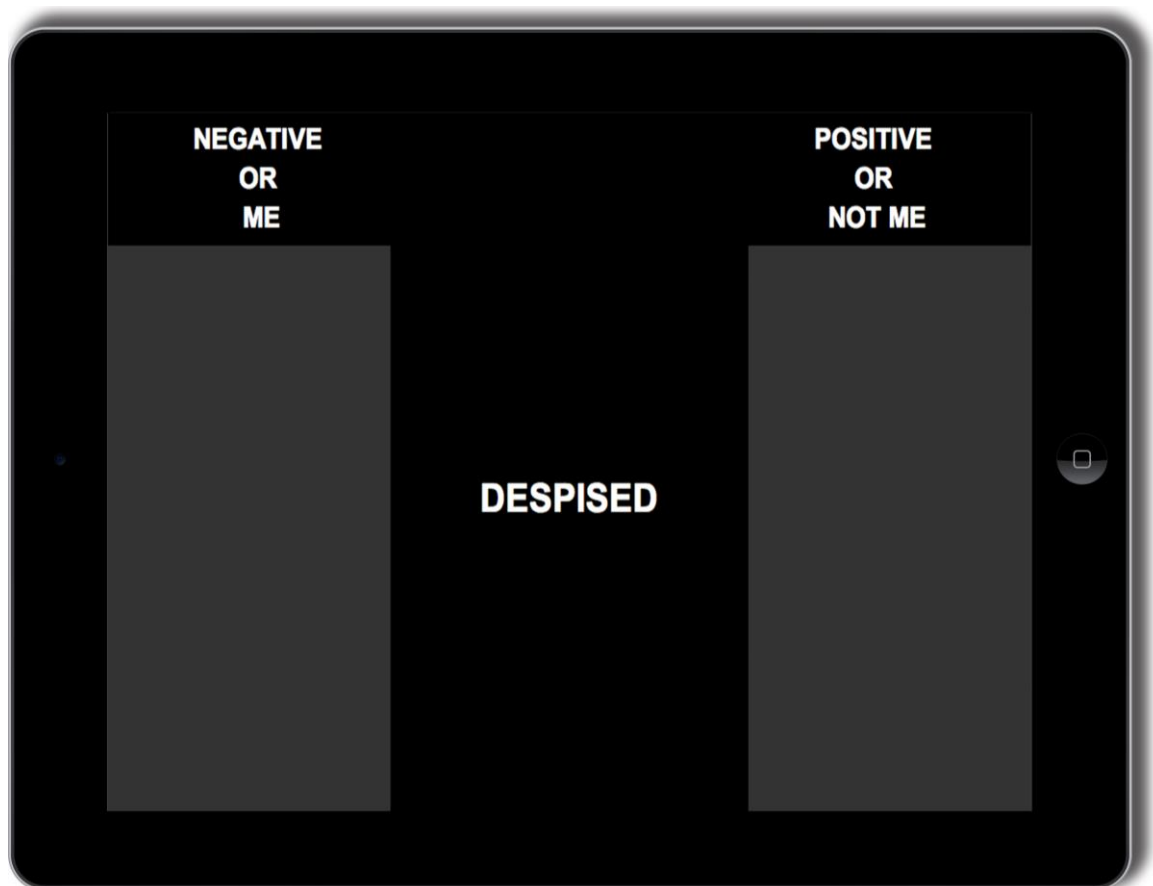
Figure 10: SE-IAT trial layout taken from Step 5. This is the way the test looks for participants. In this case, taping the left of the screen is the correct response.

Participants were counterbalanced to two different sequences of the SE-IAT: half of participants completed a Step sequence of 1-2-3-4-5, and the other half completed a sequence of 1-2-5-4-3. Implicit self-esteem is computed as the transformed difference in latency times between the compatible step (Step 3 in Table 14) and the incompatible step (Step 5 in Table 14). To calculate the final score for implicit self-esteem, or *D* score, the software was capable of recording, besides time latencies: counterbalancing information, correctness of response, trial number and the specific word which appeared on each trial.

Data Analysis

The SE-IAT was scored according to the algorithms suggested by Greenwald, Nosek and Banaji (2003).

Given that the scores yielded by the IAT are difference scores, and that the test presents the stimuli to participants in a random order, several methods to calculate internal reliability were used over difference scores. These difference scores were

calculated by subtracting the latency of a compatible trial to the latency of the corresponding incompatible trial, in order of presentation. For example, the first trial of block 5.1 minus the first trial of block 3.1, then the second trials of both blocks, and so on until 60 difference scores that were used as "items" were obtained (Bosson et al., 2000; Cunningham et al., 2001; Greenwald & Nosek, 2001; Victoria & Fontenelle, 2011). Internal consistency statistics were then obtained by different combinations of these difference scores, namely Cronbach alphas for the short (practice) blocks (3.1 and 5.1), alphas for the long blocks (3.2 and 5.2), alphas for all difference scores, split-half reliabilities of odd and even trials (in order of presentation) (Schnabel et al., 2008), alphas for the first and last 15 difference scores (in order of presentation), and alphas for the 30 difference scores presented in the middle of the test (Banse et al., 2001; Nosek et al., 2005; Schmukle & Egloff, 2004).

Differences in implicit self-esteem according to gender were evaluated by a Mann-Whitney U test for independent samples on IAT scores.

Relationships between age and implicit self-esteem were estimated using Spearman correlation coefficients and through visual inspection of data. Regression models were created in order to predict implicit self-esteem from age. Participants were separated into two groups according to age (older or younger than 18 years old), and the SE-IAT scores' differences between these groups were tested using a Mann-Whitney U test for independent samples.

All analyses were carried out using SPSS 22 on a Macintosh computer.

Results

Sample

The final sample for this study is described in Table 15. Age of participants at the time of measurement showed a normal distribution (skewness: -0.095, SE = 0.223, z = -0.43; kurtosis; -0.875, SE = 0.442, z = -1.98; S-W$_{118}$ = 0.967; all $p \leq 0.01$)

Scoring the IAT:

The first step in scoring the IAT and obtaining *D* scores is to reverse the counterbalanced participants, so the first 60 combined trials of all participants correspond to the compatible condition (blocks 3 and 4), and the last 60 combined trials represent the incompatible condition (blocks 6 and 7).

Following the improved scoring algorithm suggested by the developers of the test (Greenwald et al., 2003), the ensuing procedure has been used to obtain the final IAT *D* score:

All trials from Steps 3 and 5 have been used, which means that every participant had 60 trials in the compatible condition and 60 in the incompatible condition. However, individual trials with latencies >10,000 ms have been eliminated. In this case, of a total of 14,160 trials, 6 had to be eliminated. Cases who present more than 10% of responses faster than 300 ms must be eliminated. Only one case was eliminated on these grounds.

For each block, the mean for correct responses was calculated, as well as two pooled standard deviations for all trials (correct and incorrect): one for blocks 3.1 and 5.1, another for blocks 3.2 and 5.2. Each error latency was replaced by the latency mean of correct responses for its corresponding block plus 600 ms. Next, the mean for each block was calculated, including the corrected latency values. Subsequently, two differences were calculated: Block 5.1 − Block 3.1, which represent the difference between the compatible and incompatible conditions of the first 20 trials of each condition. The other difference is Block 5.2 − Block 3.2, which represents the difference between compatible and incompatible conditions for the last 40 trials of each condition.

	n	Age		
		Mean (SD)	Minimum	Maximum
Female	60	19.46 (2.86)	14.39	25.00
Male	58	18.92 (2.77)	14.28	24.20
Total	**118**	**19.20 (2.82)**	**14.28**	**25.00**

Table 15: Descriptive statistics for the final sample

The results of these subtractions were divided by their corresponding pooled standard deviation. Finally, both quotients were averaged, obtaining D^{7}.

D scores did not present a normal distribution. Non-parametric tests will be used when necessary. Females showed a slightly lower mean self-esteem than males (-0.039, SD = 0.60; 0.039, SD = 0.67, respectively).

Internal Consistency Analyses

Several methods to calculate internal consistency were used. They all yielded good levels of internal consistency for the SE-IAT used in this study. Reliability statistics are showed in Table 16 according to the method used.

Method	Statistic	Value (n)
All difference scores from short blocks	Cronbach's α	0.797 (20)
All difference scores from long blocks	Cronbach's α	0.803 (40)
All difference scores	Cronbach's α	0.873 (60)
Split-half reliability on difference scores of odd trials against even trials	Spearman-Brown coefficient	0.901 (60)
Difference scores of the first 30 trials and the last 30 trials	Cronbach's α	0.802 (30)
Difference scores for the 60 trials presented in the middle of the test	Cronbach's α	0.753 (30)

n = number of difference scores included in the analysis of a total of 60

Table 16: Internal reliability statistics for the SE-IAT

---

[7] SPSS syntax to obtain the *D* score is provided in Appendix I.

Relationships with Age and Gender

A Mann-Whitney U test was run to determine if there were differences in D scores between males and females. Distributions of D scores for males and females were similar, as assessed by visual inspection. Median D score was not statistically significant between genders, $U = 1.87$, $z = 0.680$, $p = 0.491$. Neither was there a correlation between age and D scores ($rho = 0.027$, $p = 0.773$). Linear and curved predictive models failed to achieve significance (linear: $R^2 = 0.001$, $p = 0.765$; quadratic: $R^2 = 0.001$, $p = 0.922$), which was noticeable during visual exploration of data, as shown in Figure 11. The group of participants who were younger than 18 years old did not significantly differ from those older than 18 years in their D scores, $U = 1,469.00$, $z = -0.517$, $p = 0.605$.
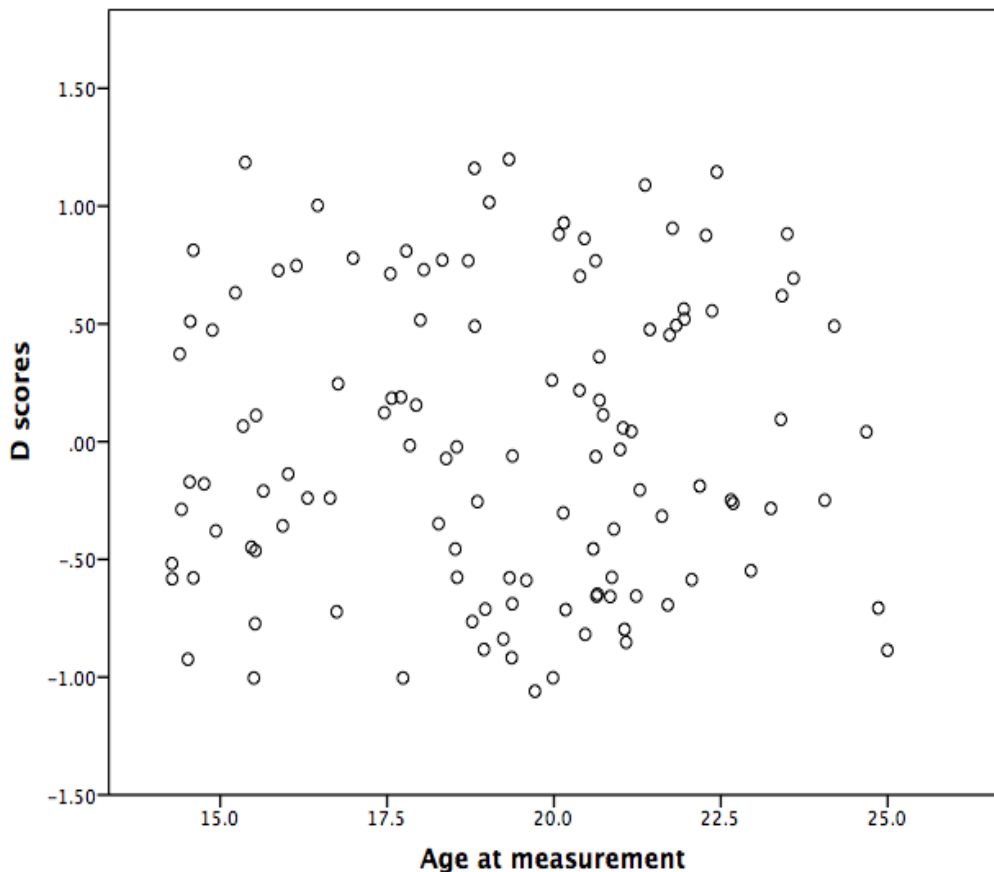

Figure 11: Scatter plot of D scores on age. The plot shows no relationship between these two variables.

Discussion

This chapter presented the SE-IAT that will be used throughout the rest of this thesis. It is an adaptation of the desktop SE-IAT to a tablet format. Various methods of calculating internal consistency showed good indexes, similar to the test's desktop version used in the literature (Nosek et al., 2007a). This is, to our knowledge, the first study to use different methods to calculate internal consistency in the IAT. These methods did yield different values, however these differences are minimal, as surmised by Schnabel (2008, p. 18). In spite of these minimal differences, the most robust calculation procedure found the highest internal consistency (0.901). As seen in the introduction to this chapter, this calculation method uses all difference scores in an odd/even presentation order to calculate split-half reliability estimates, which are robust to changes in response pattern during the test. This is because utilising difference scores between pairs of odds/even trials controls for the changes in latency that might occur as the measure approaches the end.

This version of the SE-IAT was also robust to age and gender in a normative adolescent sample, which demonstrated its discriminant validity to these demographic factors. On average, females showed slightly lower self-esteem than males, but that difference was not statistically significant. Besides lending psychometric support to this version of the measure, the lack of relationship between implicit self-esteem and age contributes to the idea of a critical period for the development of implicit self-esteem that is previous to the adolescent stage (DeHart et al., 2006; DeHart et al., 2013).

It is thus safe to conclude that the present version of the SE-IAT is a reliable, consistent and robust test to be used within the same contexts in which the desktop version has hitherto been employed. Furthermore, a portable version of this test makes it more practical for utilisation at sites where there are no desktop computers available, such as consulting rooms that are used by multiple clinicians (as it is common in public mental health centres and hospitals and many other clinical sites), allowing the test to be taken out of the lab. This is also advantageous for research contexts, especially in situations when participants are elusive or lack motivation to attend laboratories or other fixed assessment locations. In conjunction with an online storage system like POD, it concurrently allows the portability of the measure while safely storing confidential data centrally, which can be accessed instantaneously by authorised researchers in different parts of the world. Therefore, this version of the SE-IAT can be confidently used in upcoming studies in which researchers can take advantage of this measure's ease of use,

short duration and portability, which is in sharp contrast with the high-cost desktop versions (Bosson, 2006).

Limitations

It is important to note that all participants in this study completed the SE-IAT using the same tablet make, model and operative system (iPad 2 on an OSX platform). Given that the test has been built as an HTML applet tied to the device's clock, the use of different devices might affect the results when comparing different participants. It is thus imperative (at least until research is carried out to determine the real effect of the device on D scores) that the same make/model of the device is used by all participants.

This study is limited in the sense that it avoids relating D scores to other psychological constructs, like explicit self-esteem or psychopathology. However, such analyses might have been rendered useless if the measure had not shown acceptable psychometric properties.

The sample we have used in this chapter only completed the test once. This has precluded us from calculating the test-retest reliabilities for this version of the IAT. However, Chapter 6 provides a sample to which this same SE-IAT has been administered three times, with an average time span of 8 weeks between administrations. Further comments on the test-retest reliability for this SE-IAT are included there.

In the context of this doctoral thesis, it is important to remember that the results obtained through this process of validation should only be generalised with extreme care. Although it was found that this novel version of the SE-IAT possess excellent psychometric qualities, they are only warranted for a sample of normative adolescents, which diminishes its external validity, especially for generalization to the following studies. In spite of that, it was found that the IAT is consistent regarding age in a very age-heterogeneous sample, and that it is robust regarding gender. Because generalisation is not completely warranted, both the demographic variables and psychometric qualities of the SE-IAT used in the rest of the present doctoral thesis, will be calculated for each study, using the detailed knowledge on discrepancy score test validation that has been obtained in this chapter.

Chapter 4: The Relationship between Implicit Self-Esteem And Depression

Introduction

This and the following chapters will report on different aspects of the functioning of the Implicit Associations Test in depressed samples. In particular, this chapter aims at assessing the ability of the Self-Esteem Implicit Associations Test (SE-IAT) for iPad to discriminate between depressed participants and a non-depressed control group.

In Chapter 1, I have reviewed the literature covering the performance of the IAT in various areas, including psychopathology, and its use in samples with schizophrenia, anxiety disorders, eating disorders, substance use disorders, bipolar disorder, psychogenic seizures, personality disorders, and internalising problems including depression. The next section of this chapter aims to deepen on the existent literature about the use of implicit measures in depression, in order to understand the need for further studies addressing the effects of implicit cognition in this highly prevalent disorder.

Depression

Major Depression Disorder (MDD) is a heterogeneous disorder, with a highly variable course, inconsistent response to treatment, no established mechanism, and which usually presents itself in comorbidity with various psychiatric and somatic disorders. It has a worldwide lifetime prevalence as high as 20% (Belmaker & Agam, 2008; Kessler, Chiu, Demler, & Walters, 2005). It is defined as a period of at least two weeks of sustained depressed mood and/or anhedonia, accompanied by a series of psychological, vegetative and physical symptoms, such as concentration and memory difficulties, weight gain/loss and sleeping problems (American Psychiatric Association, 2010). The diagnosis carries a risk for suicide, estimated to be 3.5% (Blair-West, Cantor, Mellsop, & Eyeson-Annan, 1999; Blair- West, Mellsop, & Eyeson- Annan, 1997). Depression is a common and complex disorder that usually manifests early in life. At least 40% of depressed patients experienced their first depressive episode by age 20, and the peak annual prevalence of the disorder occurs in people aged 15 to 25 (Patten et al., 2006). The risk factors for first lifetime onset of depression are female gender, the presence of one or two short alleles of the 5HTT polymorphism, and prior alcohol abuse, drug abuse or panic attacks (Eaton et al., 2008). The hereditability rates of MDD have been shown to be around 37% in various twin studies (Belmaker &

Agam, 2008). Research shows that MDD is a recurrent disorder, with a 15-year recurrence of up to 85% in specialised mental healthcare settings and 35% in the general population, regardless of gender, socioeconomic status and civil status (Hardeveld, Spijker, De Graaf, Nolen, & Beekman, 2010). The most important predictor of recurrence is past psychiatric symptomatology, especially depressive (Caspi et al., 2014; Hardeveld et al., 2010; Kendler, Gardner, & Prescott, 2014a, 2014b). Given that depression is ranked third in terms of disease burden among all health conditions and first among all psychiatric disorders in terms of disability adjusted life years, contributions to the knowledge addressing disease mechanisms are encouraged (Everyday Health, 2013; Wittchen et al., 2011).

Neuroscientific studies have shown differences between people with MDD and non-depressed people. Structurally, a Magnetic Resonance Imaging (MRI) meta-analysis showed that depression is characterised by reduced brain volume in emotional processing areas, including the frontal, orbitofrontal and cingulate cortices, hippocampus and striatum. Depressed patients also show enlargement of the pituitary gland and excess of white matter hyperintensity lesions (known as leukoaraiosis, a series of non-specific changes in white matter which represents a risk factor for subcortical dementia) (Arnone, McIntosh, Ebmeier, Munafo, & Anderson, 2012; Sacher et al., 2012). Results regarding brain functioning in depression are mixed, depending on the imaging methodology employed. However, the overlap between different methodologies shows a modulating effect of emotional valence on functional abnormalities: while the amygdala, striatum, parahippocampal, cerebellar, fusiform and anterior cingulate cortices display hyperactivation for emotionally negative stimuli and hypoactivation for positive stimuli, the left dorsolateral prefrontal cortex shows reduced activity for negative stimuli. Furthermore, the activity of the orbitofrontal cortex in depressed patients is increased for positive stimuli (Diener et al., 2012; Fitzgerald, Laird, Maller, & Daskalakis, 2008; Groenewold, Opmeer, de Jonge, Aleman, & Costafreda, 2013; Hamilton et al., 2012). These results support the idea of a dysregulation of emotional processes, characterised by heightened automatic emotional reactions to negative environmental stimuli. These reactions are heightened to the extent that cortical, executive and explicit regulatory processes are overwhelmed, giving rise to negative mood and depressive behaviour.

From a cognitive viewpoint, depression occurs when negative self-beliefs and other processing biases hinder the capacity of an individual to regulate their emotional responses to life's adverse experiences. These negative self-biases share common

themes of loss, failure, worthlessness, rejection and hopelessness (Phillips, Hine, & Thorsteinsson, 2010). Within this theoretical frame, negative self-schemas are activated by an environmental trigger, in turn generating automatic and systematic biases in the processing of information (Hertel, 2002). The pattern of these activations is thought to reflect existing cognitive structures within an associative network, where one node of information is capable of activating other nodes in the same network. These automatic activations seem to circumvent and, in the worst cases, to recruit reflexive processes, which in non-depressed people would have modulated the automatic negative responses.(Beck, 2008).

Self-Esteem in Depression

One of the most studied psychological features in relation to depression is self-esteem. Clinicians and theorists (and common sense) have for a long time agreed that depression seems to be associated to lower self-esteem, a sense of failure, and self-dislike (Beck & Alford, 1967). Operational definitions of depression have low self-esteem as a possible symptom (American Psychiatric Association, 2010).

In spite of the scientific certainty of the strong relationship between low self-esteem and depression, the nature of this relationship is a matter for ongoing debate. While some theorists discuss self-esteem as a symptom of depression, others suggest that low self-esteem could be causal in certain types of depression (Beck, 2008; Pyszczynski & Greenberg, 1987). Recent studies and meta-analyses lend more support to the latter notion that self-esteem plays a causal role in depression. This notion is known as the vulnerability model of self-esteem and depression. The opposite notion i.e. that low self-esteem is a consequence of depression, known as the scar model, has received only limited empirical support. The vulnerability effect of self-esteem has proven to be twice as large as the scar effect (β= -0.16 vs. β= -0.08), in a large meta-analysis of 77 studies (Sowislo & Orth, 2013). The vulnerability model is equally observable in both males and females (regardless of their different levels of self-esteem and depression), and it is robust to the type of instrument used to measure self-esteem and depression (and to the common overlap between these measures), the sub-type of depressive symptoms (somatic or affective-cognitive), the time intervals between measurements (ranging from some weeks to decades), and cultural diversity between samples (Orth & Robins, 2013; Orth et al., 2012; Sowislo & Orth, 2013). Causal models that link stressful life events to the relationship between self-esteem and depression

have found that low self-esteem predicts depression even when controlling for stressful life events. It seems that it is not the presence of stressful life events which has a causal contribution to depression, but the tendency to ruminate about them which moderates the relationship between low self-esteem and depressive symptoms. Furthermore, a tendency to ruminate is more characteristic of subtypes of depression that show high levels of anxiety (Kuster, Orth, & Meier, 2012; Nolen-Hoeksema, 2000; Orth & Robins, 2013; Steinberg, 2006).

In the previous chapter we have explored self-esteem as both a general and stable trait and also a domain-specific one. However, it has been shown that it is the global (low) level of self-esteem that predicts depressive symptomatology and not its variability. That is to say, stability and contingency of self-esteem play no role in the prediction of depressive symptomatology, or that the effect of self-esteem fluctuations on depression disappears when controlling for self-esteem level. (Metalsky, Joiner, Hardin, & Abramson, 1993; Sowislo, Orth and Meier in Orth & Robins, 2013; Orth, Robins, Meier, & Conger, 2015; Wouters et al., 2013).

Considering this strong association between self-esteem and depression, it is valid to enquire about the appropriateness of conceptualising mood and self-esteem as distinct constructs in the empirical realm. Such distinction is not clear in the literature, and it appears to rely heavily on the way both mood and self-esteem are measured. For example, the Beck Depression Inventory (BDI) correlates strongly with the State Self-Esteem Scale (SEES; r= -0.71, $p<$ 0.05) (Heatherton & Polivy, 1991), and with the Rosenberg Self-Esteem Scale (RES, r= -0.60, $p<0.001$) (Osman et al., 1997). In adolescent samples, the BDI also correlates significantly with self-esteem, as measured by the RES (r= -0.42; $p=0.00$) (Harrison, 2014), and it is a powerful predictor of depression, again measured by the BDI ($\beta=-0.39$, $p<0.001$) (Cheng & Furnham, 2003). It is clear from the literature that the BDI is consistently correlated with measures of self-esteem. This relationship, however, is less clear when we consider the association between depression and implicit self-esteem being one of the features of implicit cognition that may play an important role in this disease's mechanism. Investigators have been and are interested in arriving at an understanding of the part played by implicit cognition in this disorder through different methodologies, and addressing various relevant implicit cognitive processes.

While the importance of self-esteem and other self-views is central to depression, there is a lack of theoretical consensus regarding how implicit and explicit self-esteem

exert their influence on the risk, onset, duration and relapse of depression. While certain authors posit that risk of depression depends on conscious negative self-attitudes, self-defeating reasoning and thinking styles (Nolen-Hoeksema, 2000), other researchers consider automatic processes, often preconscious, to be at the basis of the disorder (Scher, Ingram, & Segal, 2005). Although both views are backed by a plethora of research, more convincing arguments are given in favour of a dual-processing model (Chapter 1), with both implicit/automatic cognition (which requires little cognitive effort) and explicit (deliberate, effortful) cognition playing a part in depression. Negative mood is thus seen as the result of the interaction between implicit and explicit cognition, in that environmentally elicited automatic negative self-thoughts are unsuccessfully down-regulated by the more rational and effortful explicit cognitive control. The question remains, whether depression implies an augmented intensity and frequency of implicit negative self-associations that override effortful control, i.e. an accentuated bottom-up processing (Steinberg et al., 2007), or a hindrance in effortful systems to control naturally occurring implicit negative thoughts, that is to say, a dysregulated top-down processing (Haeffel et al., 2007). Functional Magnetic Resonance Imaging (fMRI) studies have found that depressed patients show both a hyper reactivity of the amygdala and the anomalous recruitment of prefrontal areas when confronted with tasks that imply the appraisal and re-appraisal of negative self-referential information (Johnstone, van Reekum, Urry, Kalin, & Davidson, 2007; Siegle, Thompson, Carter, Steinhauer, & Thase, 2007). It is probable that both explicit and implicit cognition further interact with historical risk and current environmental stressors, and that each one of these elements contributes differentially to the onset, maintenance and recurrence of depression in each person (Phillips, Hine, & Bhullar, 2012).

Implicit Measures in Depression

In order to measure implicit cognition in its relationship with depression, different measures have been employed, assessing various cognitive domains that are thought to contribute to this disorder through facilitating negative self-referential processes. Among these cognitive domains are attention, memory, interpretation, self-beliefs and, of course, self-esteem (Phillips et al., 2010).

Automatic attention is greatly influenced by previous experience and it reflects an individual's goals, emotions and moods. Depression research in implicit attention has

found anomalies in three subsystems of attention: shifting, engagement and disengagement. Biases on attention imply a facilitated shifting towards, a greater engagement with, and a disrupted ability to disengage from negative information. Among the implicit measures for attentional biases are word probe tasks: computer-based tests during which participants are presented with two words, one neutral and another which is emotionally negative. When a dot or a line replaces one of the words, subjects must rapidly react by pressing a key. A negative bias is assumed when reaction times (RTs) are shorter for negative words, implying a facilitated attentional shift to negative information (McCabe & Gotlib, 1995; McCabe & Toman, 2000). Another measure is the Dichotic Listening Task (Bruder, 1983), in which participants must concentrate on neutral words spoken in one ear, while ignoring emotionally negative words spoken in the other ear. A negative bias is assumed when the participant finds it particularly difficult to ignore negative words. Other measures vary in the emotional valence of stimuli and in the duration of the interstitial period between stimuli: this allows for researchers to the length of time that depressed patients take to disengage from the previous stimulus, when it is a negative word. Attentional disengagement effects suggest deficits in inhibitory systems: non-dysphoric participants are slower to identify the valence of both positive and negative words following primes with the same valence, while dysphoric participants show the same effect for positive valence only. That is to say, words with a negative affective valence are automatically recognised while effortful inhibition is ineffective (Joormann, 2004).

Regarding memory, non-depressed people show a bias in recalling positive information about the self, while depressed individuals exhibit biases towards negative information and away from positive self-information (Matt, Vázquez, & Campbell, 1992). The most common method used to measure the implicit effects of memory is priming, where participants' responses indicate exposure to previous material. In the Lexical Decision task (Clark, Teasdale, Broadbent, & Martin, 1983) lexical stimuli are briefly presented, masked, and then presented again. These stimuli comprise words and non-words (a meaningless string of letters that resemble a word). On the second presentation of the stimulus, participants must rapidly recognise if the displayed string of letters is a valid word or not. A negative bias is assumed when RTs are faster for negative words. Literature reviews on the topic indicate that results in depression are inconsistent, with some studies finding differences between depressed and non-depressed groups, while others do not (Gotlib & Joormann, 2010; Wisco, 2009).

Interpretation of life events and self-beliefs are also characteristically altered in depression. In Wisco's review (2009) these include beliefs about personal inadequacy, perfectionistic self-standards, self-blame for real or imagined bad outcomes, expectation of rejection by others, and pessimistic views about one's own future. However, only a few methods have been developed to assess these features in their implicit form. The most common one is the Scrambled Sentences Task (SST) (Rude, Wenzlaff, Gibbs, Vane, & Whitney, 2002; Wenzlaff & Bates, 1998) in which the participant is presented with ambiguous information (e.g. "looks the future bright very dismal"), and are asked to use 5 of the 6 words to create the first correct sentence that comes to mind. The task is completed under time pressure and the purpose of the task is obscured by presenting neutral sentences. An implicit depressive bias is assumed when the person creates a sentence with negative content ("the future looks very dismal") instead of positive ("the future looks very bright"). Other methods follow the same logic: asking participants to write down homophones presented verbally. These could be interpreted either positively or negatively (e.g. negative: "mourning-morning"; positive: "peace-piece") (Wenzlaff & Eisenberg, 2002). Other studies finding associations between implicit measures and depression, manipulate participants' cognitive load or distraction, which increases the depressive associations, giving further support to the idea of an automatic processing of negative stimuli which is only partially modulated by reflexive processes (Rude et al., 2002; Wenzlaff & Bates, 1998; Wenzlaff & Eisenberg, 2002).

To date, results of these measures in depression are equivocal. Researchers who have obtained meaningful results in depression tend to be those who use self-related stimuli (Gotlib & Joormann, 2010; Wisco, 2009). This highlights the importance of self-directed negativity in depression as distinctive, in contrast to a more general negativity. In this same line, beliefs about one's own future have been found to be present as automatically negative predictions in depressed patients. The use of ideographic and self-referential material in the study of implicit cognition on depression is, in one way or another, measuring the valence of attitudes towards the self, self-esteem being among them.

Implicit self-esteem has not only been studied in relationship to the differences between people who are depressed and those who are not, but in prospective designs, trying to establish if implicit self-esteem is a protective or risk factor for the future development of depressive symptomatology. For example, the Name Letter Preference Task (NLPT, as seen in Chapter 1) has been utilised to predict depression in a 6-month longitudinal study (Franck, Raedt, & Houwer, 2007b). Formerly depressed patients,

currently depressed patients and never-depressed controls showed equal levels of implicit self-esteem, while currently depressed patients showed a significantly worse explicit self-esteem than the other groups. However, when controlling for initial depressive symptoms, only low implicit self-esteem was able to predict depressive symptoms after 6 months for all groups. As with other implicit cognition tasks on depression, results are equivocal. A 3-month follow-up study using the NLPT found that depressive symptomatology was predicted only by the discrepancy between high explicit self-esteem and low implicit self-esteem, moderated by conscious rumination of negative thoughts (Phillips & Hine, 2014).

Another measure of implicit self-esteem, the Extrinsic Affective Simon Task (EAST) was used in depressive patients (De Raedt, Schacht, Franck, & De Houwer, 2006). This is a reaction time computerised test, very similar to the IAT, but without the contrasting categories "self-other": presented word stimuli were either self-referential or not, and either positive or negative. Participants must categorise words depending on their colour. However, valence of the word and self-referential content has an effect on response time. This particular experiment showed that depressed patients' responses to coloured self-items were faster with self-referential positive words, giving an indication of statistically significant higher implicit self-esteem than controls.

As we can see from this revision of the literature, results are contradictory. It is certainly possible that different measures address distinct underlying constructs. It might be that focusing on a single measure, in this case the SE-IAT, will yield a more consistent pattern of results in depressed patients.

The IAT in depression. The Implicit Associations Test has been extensively utilised in depressed samples (Remue, Hughes, De Houwer, & De Raedt, 2014). It is possible to find different versions of the test measuring various psychological constructs believed to contribute to this disease's risk, markers, and mechanisms.

For example, a Depression-IAT has been developed, where the target concepts "self" and "other" are paired to attributes related to "depressed" (e.g. useless, pessimistic, inadequate) and "elated" (e.g. positive, optimistic, active). Participants were also asked, after taking the IAT, to answer a self-report where they had to score each one of the "depressed-elated" words from 1 to 5 depending on "to what extent you think it generally applies to you (Elgersma, Glashouwer, Bockting, Penninx, & de Jong, 2013, p. 954; Glashouwer & de Jong, 2009, p. 1103)". Glashouwer et al. measured depressed patients, anxious patients, comorbid anxious-depressed patients and controls. They

found that depressed patients showed stronger depressive self-associations than anxious patients and controls, but comorbid patients tended to show even stronger depressive self-associations. Further analyses showed that remitted depressive patients also presented stronger self-depressive associations than controls, but significantly weaker associations than currently depressed patients. Implicit depressive self-associations predicted depressive symptom severity over and above explicit self-associations ($F_{1, 2827} = 18.12$, $p < 0.001$, R^2 change=0.003). A similar pattern was found regarding explicit self-associations. However, depressive self-associations were unrelated to anxiety (Glashouwer, de Jong, & Penninx, 2011). Elgersma et al. (2013) found that explicit and implicit depressive self-associations were not related to each other, but that both implicit and explicit negative self-associations correlated positively with both the number of previous depressive episodes and with the duration of depressive symptoms between baseline and 2-year follow-up. The latter was a much stronger relationship: the relation between implicit self-esteem and previous depressive episodes fell barely out of significance when neuroticism was entered as covariate (number of previous episodes: $\beta = -0.07$, $p = 0.06$). This result gives further support to a vulnerability model of self-esteem and depression in contrast with the scar model, and it is relevant regarding assessment of risk for future depressive episodes. A similar study with the same IAT, compared controls and depressive patients with bipolar patients. These latter two groups showed significantly stronger depressive self-associations than controls. Explicit and implicit depressive self-associations were related to each other only in unipolar depressive patients (Jabben et al., 2014).

As we can see, results with the Self-Referential Depressive IAT are more consistent throughout the literature, where depressed patients show similarly strong implicit self-depressive attitudes (notwithstanding that some studies find relationships between implicit and explicit self-esteem and others do not). In the case of the IAT measuring implicit self-esteem, this consistency is no longer present.

The SE-IAT in depression. Several experimental studies have used the SE-IAT in relation to depression. Their results are mixed, mostly classifiable into three groups: those that find the expectable lower implicit self-esteem in depressed patients compared to non-depressed, those that, contrary to expectations, find that depressive patients show higher self-esteem than controls, and those that find no relationship between implicit self-esteem measured with the SE-IAT and depression. Pertaining to the first group, Conner and Barrett (2005), in a normative student sample, found that stronger associations of the self with unpleasant in contrast to pleasant (e.g. peace, sunrise vs.

death, vomit, etc.) are able to predict levels of boredom, disappointment, shame, anger, disgust, and other negative emotions in the following 26 days (spontaneous affect was reported by subjects 10 times a day). This predictive power remained statistically significant only for anger and disgust when explicit measures of self-esteem were entered as covariates. This implies an overlapping effect of implicit and explicit measures (their correlation coefficient was $r= 0.23$, $p<0.001$). A second study reported in the same paper (Conner & Barrett, 2005, p. 479), used a different IAT in which the categories pleasant/unpleasant were shifted to categories of success vs. failure (words like accomplish and winner vs. ashamed and incapable). As with the previous experiment, participants received an electronic personal digital assistant (PDA), which would remind them to complete a 16-item questionnaire ten times a day, each item addressing a different type of affect-related experience for 17 days. In this case, the predictive effect of implicit self-attitudes remained significant for negative events after the inclusion of explicit measures in the analyses. A similar study found that the predictive power of implicit self-esteem on depression incidence is lost when the baseline depressive symptoms, cognitive reactivity, and number of negative life events were entered as covariates (Kruijt et al., 2013).

The second group of studies finds that a higher level of implicit self-esteem is related to depressive manifestations. These results are interesting, when considering that the assumed relationship between self-esteem and depression is an inverse one. Focusing on adolescent offspring (aged 13-19 years old) of bipolar patients, a study found that a pattern of high implicit self-esteem and low explicit self-esteem was associated with depressive features, while the opposite self-esteem configuration was associated to features of mania, and explicit and implicit self-esteem were not associated to each other (Pavlickova, Turnbull, & Bentall, 2014). Similar high levels of implicit self-esteem were also found not only in depression, but also in patients with social phobia and body dysmorphic disorder (Roefs et al., 2011).

The third group of results attempts to relate depression with implicit self-esteem as measured by the SE-IAT. In these studies, implicit self-esteem IAT scores are not able to differentiate between depressive and non-depressive participants and, in general, are not able to predict depressive symptomatology (De Raedt et al., 2006; Franck, De Raedt, & De Houwer, 2008). Kesting, Mehl, Rief et al. (2011) compared schizophrenic patients with acute persecutory delusions (n= 28), those with remitted persecutory delusions (n= 31), depressed patients (n= 21), and healthy controls (n= 59). They found

that both schizophrenic and depressed patients did not differ from the control group in their levels of implicit self-esteem.

A study mentioned above (De Raedt et al., 2006) concluded that depressed patients had high implicit self-esteem by comparing reaction times of the IAT's compatible block against reaction times in the incompatible block. However, they found the same pattern in the non-depressed group: high implicit self-esteem was a characteristic shared by the whole sample. A similar study by the same research group replicated the null finding, but only between depressive patients with suicidal ideation and healthy controls, while depressed patients without suicidal ideation showed significantly lower implicit self-esteem than the control group (Franck, De Raedt, Dereu, & Van den Abbeele, 2007a). A further replication of these studies found no difference in implicit self-esteem as measured with the IAT between currently depressed individuals and controls. However, there was a significant difference between controls and individuals with a remitted depression, the latter having higher implicit self-esteem. This difference disappeared after a depressive mood induction procedure. In the first study mentioned in this paragraph, the null finding arose only when using the SE-IAT; depressive patients had higher implicit self esteem than controls when measured with the EAST (De Raedt et al., 2006). A similar study found that while implicit self-esteem measured with the IAT was unrelated to depression, when the same attitudes were measured with the NLPT, the main effect of implicit self esteem on depression was significant. The two measures of implicit self-esteem were not related to each other ($r= -.06$, $p > .05$) (Lima, 2007). Studies using adolescent samples also find no relationship between implicit self esteem and symptoms of depression and anxiety (Bos, Huijding, Muris, Vogel, & Biesheuvel, 2010; De Jong et al., 2012; van Tuijl, de Jong, Sportel, de Hullu, & Nauta, 2014).

The SE-IAT on Suicide. The IAT has also been used in the study of suicide, due to the fact that direct prompting about suicide does not always achieve disclosure. A relatively recent study showed that up to 78% of patients who die by suicide deny suicidal thoughts in their last verbal communications before dying (Busch, Fawcett, & Jacobs, 2003). A special type of IAT, the Death/Suicide IAT (D/S-IAT) attempts at measuring implicit associations between the self and either life or death. Scores on the D/S-IAT have been found to be strongly related to suicidal ideation and acts, but their predictive ability is disputed. Two studies found that the predictive power of the D/S-IAT improves the overall prediction of suicide attempts within 3 to 6 months following psychiatric emergency care (Randall, Rowe, Dong, Nock, & Colman, 2013), sometimes

even exceeding the predictive value of well-established markers, such as a history of previous suicide attempts or the presence of depression (Nock et al., 2010). These same authors measured implicit self-associations with self-injury and found that the IAT was able to predict severity of ideation and attempt status (attempters vs. non-attempters). This IAT was able to incrementally improve prediction of suicide outcomes above and beyond risk factors (Nock & Banaji, 2007). However, another study, this time with an adult community sample, showed associations between the D/S-IAT and indicators of suicide risk (ideation frequency, ideation intensity, depression, thoughts about and attempts to self-harm and their intensity), but had much weaker predictive ability (Harrison, Stritzke, Fay, Ellison, & Hudaib, 2014). Another study with this same IAT, and another measuring associations between self and escape-related words in treatment resistant depressive patients, found that the latter IAT showed weaker associations between self and escape related words after a single injection of ketamine. The D/SIAT was unrelated to all measures of depression and suicide and it did not change after the administration of ketamine (Price, Nock, Charney, & Mathew, 2009). A study with currently depressed patients, patients with remitted depression and non-depressed controls found that only currently depressed patients without suicidal ideation had significantly lower levels of implicit self-esteem. Currently depressed patients with suicidal ideation showed similar levels of implicit self-esteem as non-depressed controls (Franck et al., 2007a). While it is clear that suicide is an important topic within depression where participants may have motivations not to fully disclose, the differences between these studies probably reflects the different characteristics of the samples used in the studies, hence there is no evidence to achieve a scientific consensus on the matter.

Hypothesis

Given that the literature shows a great diversity of designs and contradictory results regarding the relationship between the SE-IAT and depression, it is impossible to specify a particular set of hypotheses at this stage. For this reason, the aim of this study is an exploratory one. Its objective is to explore the relationships of the SE-IAT with demographic characteristics of the sample and with measures of depression, and to further explore the significant relationships that might be discovered. These relationships will inform hypotheses for the next chapter.

Methods

Design

This is a controlled cross-sectional correlational study. To explore the performance of the test, the SE-IAT was administered to depressed adult patients and a non-depressed control group, together with other validated measures which tap specific symptomatic components of depression.

Procedure

Participants were invited to the Institute of Neurology at UCL to complete a battery of self-report tests, including the SE-IAT (procedure explained below) and to carry out simple neurocognitive tasks. These latter tasks were part of a larger neuroimaging study, which obtained data on participants' neurocognitive performance while being scanned with Functional Magnetic Resonance Imaging (fMRI) techniques (not reported here). The battery of self-report measures was completed before participants entered the scanner.

Depressed volunteers were invited to participate from NHS organisations (including GP practices), the community (including electronic media), and from a larger randomised controlled trial (REDIT, details given in Chapter 5).

Flyers were sent to GP practices and other NHS sites explaining the study and asking professionals to make any potentially suitable participants suffering for depression aware of the opportunity to participate in this research. A short information sheet was included for clinicians to pass to potential participants, which provided the contact details of the research team in order for the individual to contact them directly if they were interested in taking part. Additionally, there was a 'tear-off' slip for potential participants to agree to be contacted by the research team, which was to be returned via the clinician.

Adverts were placed in the community and, in particular, at self-help groups. Potential participants in the community were able to contact the research team directly through a website. These participants were asked for consent to contact their GP or mental health worker to ensure the participant met the inclusion criteria outlined below. If participants were unwilling to give these details, the diagnosis of depression was confirmed though the Structured Clinical Interview for DSM-IV, Axis I (SCID-I).

Potentially suitable patients who had been screened for REDIT were informed of the imaging study by their health provider, who obtained their consent to be contacted by researchers. If consent was given, the participant's contact details were passed to the research team, who contacted the potential participant with further information about the study. Subsequent arrangements were then made to provide potential participants with a written explanation of the study and, if applicable, a visit to the Institute of Neurology was scheduled to conduct a full screening for eligibility.

Coordination of the recruitment of non-depressed controls was carried out by the Institute of Neurology at UCL. The study was advertised to non-depressed volunteers using a poster to be displayed in public-access areas, such as GP practices' waiting rooms, community centres, and through electronic sources.

The participation of all subjects was voluntary and consensual and they were paid for their time.

Participants

To be included in the study, depressed participants had to be 18 years of age or older, right-handed, fluent in written and spoken English and to have a current diagnosis of major depressive disorder (MDD). Exclusion criteria for depressed participants were the presence of clinically significant learning difficulties, health contraindications to MRI scanning (e.g., the presence of metal in the body), current psychotic symptoms or bipolar disorder, current use of antipsychotic medication, comorbidity with complex personality disorder, historic or current self-injury/parasuicide, historic or current eating disorder, and current harmful use of drugs and or alcohol. Non-depressed controls were also 18 years or older, fluent in English, right-handed and with a score below 12 in the Patient Health Questionnaire (PHQ-9). Exclusion criteria for control participants included historic or current mental health problems of note, and contraindications to MRI scanning.

The total sample consisted of 61 participants (41 depressed patients and 20 controls). Demographic characteristics of the recruited sample are shown in Table 17.

Female (n)	Depressed	26
	Control	10
	Total	**36**
Male (n)	Depressed	15
	Control	10
	Total	**25**
Age mean (SD)	Depressed	35.29 (9.88)
	Control	33.75 (8.61)
	Total sample	**34.79 (9.43)**

Table 17: Age, gender and diagnosis of the depression of this study's sample

Measures

Hamilton Depression Rating Scale (HDRS-17) (Hamilton, 1960). A structured interview designed to quantify the severity of depressive symptoms in patients already diagnosed as suffering from depressive disorder. This measure has been used extensively in psychological research for more than 50 years with different populations, languages and cultures. The psychometric properties of the instrument range from adequate to excellent. In an extensive review covering 13 years of studies using the HDRS (Bagby, Ryder, Schuller, & Marshall, 2014), Cronbach's alphas for internal consistency ranged from $\alpha = 0.46$ to 0.97 (mean $\alpha = 0.74$, sd$= 0.14$; k=13). In this study, the HDRS showed an internal consistency of $\alpha = 0.48$. Inter-rater reliability has been reported to be very high for the HDRS-17 total score, ranging from 0.80 to 0.98 (Cusin, Yang, Yeung, & Fava, 2010). Researchers using the instrument in the present study engaged in a one hour per week peer supervision to improve inter-rater agreement as the research group was already using this measure in other trials. Cut-off points for this measure yield four groups: not depressed, and mild, moderate, and severely depressed.

The 14-Item Snaith–Hamilton Pleasure Scale (SHAPS) (Snaith et al., 1995). A self-report questionnaire used to measure hedonic capacity. Each of the items has a set of four response categories: Definitely Agree, Agree, Disagree, and Strongly Disagree, with either of the Disagree responses receiving a score of 1, and either of the Agree responses receiving a score of 0. Thus, the SHAPS final score is obtained by adding the

scores from 14 items so that total scores ranged from 0 to 14. A higher total SHAPS score indicated higher levels of a present state of anhedonia. The SHAPS' original internal consistency was 0.86, as estimated by the Kuder-Richardson formula for non-parametric data (Snaith et al., 1995), and good indicators of validity. More recent assessment of the SHAPS has revealed a Cronbach alpha of 0.91 (Nakonezny, Carmody, Morris, Kurian, & Trivedi, 2010).

Beck Depression Inventory, Second Edition (BDI-II). This is a 21-item self-report inventory assessing depressive symptomatology (Beck, Steer, & Brown, 1996). Each of the items is rated on a 0-3 scale with summary scores ranging between 0 and 63. It has been found to have high internal consistency with alphas of 0.91: 0.93 among college students and 0.92 among outpatients (Dozois, Dobson, & Ahnberg, 1998). In this sample, the BDI showed a Cronbach's $\alpha = 0.96$.

Patient Health Questionnaire, Depression Module (PHQ-9). This is a 9-item self-administered questionnaire. It is the depression module of a longer test, the PHQ, which covers 8 DSM-IV diagnoses. Responses to the items (each representing one DSM criteria) range from 0 ("not at all") to 3 ("nearly every day"). Major depression is diagnosed if 5 or more of the items are present at least "more than half the days" in the past two weeks, and one of the symptoms present is depressed mood or anhedonia. The suicide item can be scored if it has been found to be present at some stage, even if this is not in the last two weeks. Internal consistency ranges from $\alpha = 0.86 - 0.89$, which is considered excellent. A cut-off point for major depression is a score of ≥ 10, with both sensitivity and specificity of 88%, and a positively likelihood ratio of 7.1 for this cut-off point (Kroenke & Spitzer, 2002; Kroenke, Spitzer, & Williams, 2001). In this study the PHQ-9 yielded an $\alpha = 0.932$.

Self-Esteem Implicit Associations Test (SE-IAT) (Greenwald et al., 1998). As we have explained in greater detail in previous chapters, the IAT is a computerised time-reaction instrument, which measures the differential implicit/automatic association of 2 target concepts with an attribute. In the case of the Self Esteem-IAT, the two target concepts are "me" vs. "not me", which are paired with the valence "good" or "bad". This study utilises the same SE-IAT described in the previous chapter, on the same hard- and software platforms, namely iPad 2 and Patient Owned Database (POD).

Data Analysis

Data was analysed using SPSS version 22 for Mac. Gender effects on diagnosis of depression were assessed using chi-squared tests. Age effects on depression diagnosis will be explored with t-tests. A zero order correlation matrix will serve to explore the relationship of implicit self-esteem and continuous symptomatic measures. If demographic influences on diagnosis are found, then they will be controlled in a logistic regression of depression diagnosis on IAT D score. Such demographic characteristics will be controlled for in subsequent multiple regressions of scores in symptomatic measures predicted by IAT D score. If no demographic differences are found, then diagnostic group differences on D score will be assessed with t tests, and continuous symptomatic measures will be the outcome of a linear regression with D scores as the only predictor. Other relevant correlations between D scores and symptomatic measures of depression will be analysed in detail using cumulative odds ordinal logistic regressions, given the ordinal nature of the symptomatic measures' subscales. Bonferroni type adjustments will be used when variables are subjected to multiple comparisons. If relevant relationships are found between both implicit and explicit self-esteem with symptomatic variables, they will be explored with hierarchical multiple regression techniques, entering explicit self-esteem variables in the first step, and implicit in the second, to ascertain the unique contribution of implicit self-esteem to the prediction of the relevant symptomatic outcome.

Results

Sample

Age approached a normal distribution throughout the sample (Shapiro-Wilk = 0.962, p = 0.053). Both depressed and not depressed participants showed no differences in terms of age (t = -0.60, p = 0.553) or gender ($\chi^2_1 = 1$, $p = 0.317$). The distribution of depression severity according the HDRS across the sample is shown in Figure 12

Total scores for all measures of depression were strongly correlated with each other, as displayed in Table 18 which denotes an adequate level of general measurement integrity.

	1	2	3	4
1. PHQ-9	1			
2. HDRS-17	0.898**	1		
3. SHAPS	0.874**	0.795**	1	
4. BDI-II	0.940**	0.905**	0.833**	1

**. Correlation is significant at the 0.01 level (2-tailed).

Table 18: Pearson correlations between total scores of depression measures. (PHQ-9: Patient Health Questionnaire, Depression Module; HDRS-17: Hamilton Depression Rating Scale; SHAPS: Snaith-Hamilton Pleasure Scale; BDI-II: Beck Depression Inventory, Second Edition)

D scores were calculated using the same procedure described in the previous chapter (Greenwald et al., 2003). The *D* score did not show normality (Shapiro-Wilk = 0.91, p = 0.000) and was negatively skewed.[8] Therefore, in order to be able to use parametric tests on *D* scores, these were reciprocally exponentially transformed and reversed. The transformed *D* score was distributed normally (Shapiro-Wilk = 0.99, p = 0.893).

Relation of the D Score with Demographic Variables.

Means for the *D* score were positive: 0.44 for depressed patients and 0.64 for controls, which were significantly different from zero (t = 16.18 and 13.90, respectively; p=0.000)
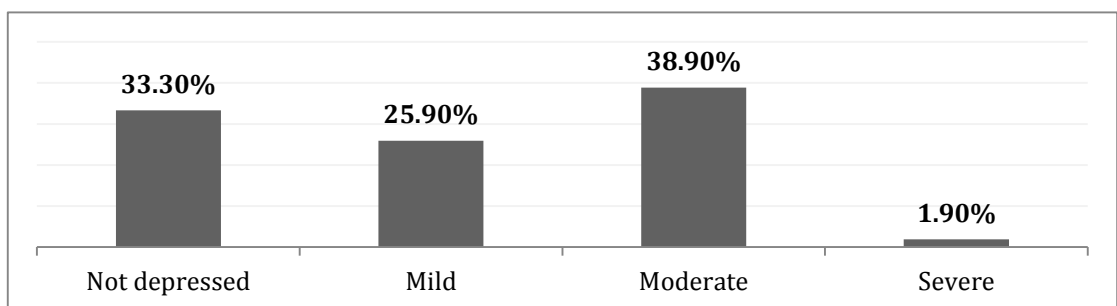

Figure 12: Severity of depression according to the HDRS-17.

---

[8] Former ways to calculate *D* scores included a logarithmic transformation of time-reaction data, which yielded a normal distribution for *D*. More modern and improved procedures do not transform data (Greenwald & Farnham, 2000; Greenwald et al., 2003).

Age showed no association to the *D* score (Pearson's *r* = -0.15, *p* = 0.91). Males and females did not show significant differences in *D* scores, tested with a *t*-test (t = -0.16, *p* = 0.88).

Implicit Self-Esteem and Measures of Depression

Diagnostic group differences in *D* scores were explored. A student's t-test did not show differences between depressed and control participants in *D* scores (*t* = -0.156, *p* = 0.877).

More minute analyses, namely a cumulative odds ordinal logistic regression, were run to determine the effect of *D* scores on severity of depression as measured by the HDRS-17. Both the deviance goodness-of-fit and the Pearson goodness-of-fit tests indicated that the model was a good fit to the observed data, $\chi^2_{(158)}$ =118.615, *p* = 0.992; $\chi^2_{(158)}$ = 176.693, *p* = 0.147. Likewise, the final model statistically significantly predicted the dependent variable over and above the intercept-only model, $\chi^2_{(1)}$ = 6.378, *p* = 0.012. An increase in a unit of *D* score was associated with a decrease in the odds of presenting more severe depression by the HDRS-17, with an odds ratio of 0.234 (95% CI, 0.070 to 0.785), Wald $\chi^2(1)$ = 5.526, *p* = 0.019.

Zero order correlations were calculated between transformed *D* scores and symptomatic measures. Correlations between *D* scores and symptomatic measures are displayed for each measure and its subscales are displayed in the tables below. Most correlations are negative, which highlights the fact that while symptomatic measures' scores increase, implicit self-esteem decreases. Correlations of *D* scores with subscales of the PHQ are displayed in Table 19. Given the multiple test the transformed *D* score was subjected to, Bonferroni corrections were used to modify the significance threshold, due to the elevated risk of Type I errors when multiple comparisons are tested. Table 19 shows significance values before the adjustment, and also shows the acceptable significance threshold after the type adjustment.

PHQ-9 Subscale	D score r (p)
Little interest or pleasure in doing things	-0.168 (0.197)
Feeling down, depressed or hopeless	-0.205 (0.112)
Trouble falling or staying asleep	-0.240 (0.062)
Feeling tired or little energy	-0.216 (0.095)
Poor appetite or overeating	-0.158 (0.223)
Feeling bad about yourself	-0.201 (0.120)
Trouble concentrating	-0.158 (0.225)
Retardation or agitation	-0.199 (0.123)
Thoughts of self-harm or suicide	-0.330 (0.009)
Total score	-0.239 (0.064)

Bonferroni-adjusted *p<0.005.*

Table 19: Pearson correlations between transformed D scores and subscales of the Patient Health Questionnaire, Depression Module (PHQ-9). Significance values for each correlation are unadjusted, but significance threshold after the Bonferroni type adjustment is showed at the bottom of the table..

After adjusting significance values using Bonferroni corrections, none of the correlations between the transformed *D* scores and subscales of the PHQ-9 were statistically significant. In spite of this, and considering both that the Bonferroni type adjustment is a conservative test (and so it increases the risk for Type II errors), and that the negative correlation found between implicit self-esteem and the frequency of thoughts about self-harm and/or suicide approaches this conservatively corrected new significance threshold, a cumulative odds ordinal logistic regression was run, with *D* scores as predicting parameters for frequency of thoughts about self-harm and suicide reported on the PHQ-9. This regression model statistically significantly predicted the dependent variable over and above the intercept-only model ($\chi^2_1 = 7.428$, $p = 0.006$). An increase in *D* scores was associated with an increase in the odds of thinking about self-harm and suicide nearly every day, with an odds ratio of 0.337 (95% CI, 0.145 to 0.780), Wald $\chi^2_1 = 6.455$, $p = 0.011$.

There was a negative correlation between the total PHQ scores and *D* scores, which only tended to significance before the Bonferroni correction, both for the transformed *D* scores (Pearson correlation, Table 19) and for the non-normal *D* scores

HDRS-17 Subscales	D score r (p)
Depressed mood	-0.173 (0.190)
Feelings of guilt	-0.092 (0.487)
Suicide ideation	-0.297 (0.022)
Early insomnia	-0.159 (0.239)
Middle insomnia	-0.209 (0.132)
Late insomnia	-0.125 (0.373)
Impairment in work and activities	-0.078 (0.549)
Retardation	-0.031 (0.814)
Agitation	-0.047 (0.725)
Anxiety - Psychic	-0.116 (0.328)
Anxiety - Somatic	0.019 (0.855)
Somatic symptoms: Gastrointestinal	-0.031 (0.815)
Somatic symptoms: General	-0.135 (0.309)
Genital symptoms	-0.139 (0.293)
Hypochondriasis	-0.024 (0.855)
Loss of weight	0.301 (0.020)
Insight	-0.018 (0.912)
Total score	-0.166 (0.201)

Bonferroni-adjusted *p<0.003.*

Table 20: Pearson correlations between transformed D scores and subscales of the HDRS-17. Please note that the significance level for rejecting the null hypothesis after a Bonferroni type adjustment is p<0.003.

(Spearman's rho= -0.247, *p*=0.055). This lack of significance must be interpreted with caution. This association might achieve significance in a larger sample.

Table 20 shows the Pearson correlation coefficients between transformed *D* scores and subscales of the HDRS-17. After correcting for multiple comparisons, using a Bonferroni correction, none of the correlations achieved significance.

Table 21 shows transformed *D* scores in relationship with subscales of the SHAPS, and with the SHAPS total score for anhedonia. For these analyses, port-hoc Bonferroni corrections were also used, rendering non-significant all the correlations.

SHAPS subscales	*D* score r (p)
I would enjoy my favourite television or radio programme	-0.127 (0.474)
I would enjoy being with my family or close friends	-0.085 (0.515)
I would find pleasure in my hobbies and pastimes	-0.307 (0.079)
I would be able to enjoy my favourite meal	-0.437 (0.010)
I would enjoy a warm bath or refreshing shower	0.108 (0.545)
I would find pleasure in the scent of flowers or the smell of a fresh sea breeze	-0.026 (0.841)
I would enjoy seeing other people's smiling faces	-0.443 (0.009)
I would enjoy looking smart when I have made an effort with my appearance	-0.096 (0.590)
I would enjoy reading a book, magazine or newspaper	-0.179 (0.311)
I would enjoy a cup of tea or coffee or my favourite drink	-0.147 (0.258)
I would find pleasure in small things	-0.266 (0.128)
I would be able to enjoy a beautiful landscape or view	-0.372 (0.030)
I would get pleasure from helping others	-0.128(0.470)
I would feel pleasure when I receive praise by other people	-0.258 (0.045)
Total score	-0.330 (0.057)

Bonferroni-corrected $p < 0.003$

Table 21: Pearson correlations between D scores and subscales of the SHAPS

Once again, the valence of most correlations are negative, implying increases in anhedonia associated with decreases in implicit self-esteem.

As shown in Table 22, *D* scores would correlate with the BDI-II's suicide

BDI-II subscales	*D* score
	r (p)
Sadness	-0.111 (0.531)
Pessimism	-0.379 (0.027)
Feeling like a failure	-0.211 (0.103)
Loss of pleasure	-0.323 (0.062)
Guilty feelings	-0.252 (0.150)
Punishment feelings	-0.038 (0.831)
Self-dislike	-0.237 (0.066)
Self-criticalness	-0.347 (0.044)
Suicidal thoughts or wishes	-0.343(0.047)
Crying	0.156 (0.380)
Agitation	-0.126 (0.333)
Loss of interest	-0.155 (0.380)
Indecisiveness	0.146 (0.410)
Worthlessness	-0.404 (0.018)
Loss of energy	0.141 (0.279)
Changes in sleep pattern	0.054 (0.763)
Irritability	0.003 (0.987)
Changes in appetite	-0.138 (0.435)
Concentration difficulties	-0.133 (0.306)
Tiredness or fatigue	-0.131 (0.460)
Loss of interest in sex	-0.127 (0.474)
Total score	-0.240 (0.172)

Bonferroni-corrected *p<0.002*

Table 22: Pearson correlations between D scores and the BDI-II

subscale, and with several subscales of self-esteem until the significance threshold is corrected with a Bonferroni type adjustment. However, given the pattern of associations between D scores and subscales of self-reported self-esteem in the BDI-II, and self-reported suicide in the PHQ-9, HDRS-17 and the BDI-II, a hierarchical multiple regression was run to determine whether the addition of implicit self-esteem improved the prediction of suicidal risk over and above self-reported self-esteem alone. In order to avoid multicollinearity issues, a new variable was created, which merges responses in the BDI-II's self-esteem items (feeling like a failure, self-dislike, self-criticalness, and worthlessness; all of them scored from 0 to 3, where higher scores represent lower self-esteem).

In order to merge these variables into one, each of the items was standardised to a Z score. Internal consistency of these 4 new standardised variables was $\alpha=0.96$. Given this high consistency, the standardised values were averaged into a variable called "BDI Self-Esteem", where higher scores represent lower self-esteem. Using the BDI-II as a proxy for self-reported (implicit) self-esteem makes sense in the light of the high correlation between these BDI items and measures of explicit self-esteem in the literature (Gorenstein, Andrade, Tung, & Artes, 1999; Heatherton & Polivy, 1991). Correlations between this composite variable of explicit self-esteem and the D score did not achieve significance ($rho= -0.132$, $p=0.312$).

Regarding suicide variables, a similar new variable was created, standardising and averaging suicide subscales of the PHQ-9, HDRS-17, and BDI-II, all of which are also scored from 0 to 3, where higher scores represent increased severity of suicide ideation. Cronbach alpha for these suicide variables was $\alpha=0.87$. See Table 23 for full details about each regression model forming part of the hierarchical process. The full model of explicit and implicit self-esteem (Model 2) was statistically significant ($R^2 = 0.322$, $F_{(2, 58)} = 13.753$, $p < 0.001$; adjusted $R^2 = 0.298$). The addition of implicit self-esteem to the prediction of suicide risk led to a statistically significant increase in R^2 of 0.053, $F_{(1, 58)} = 4.543$, $p= 0.037$. The effects of both explicit and implicit self-esteem are inverse: higher levels of both implicit and explicit self-esteem predict lower levels of suicidal ideation. However, a One Way ANOVA did not discriminate between controls, depressed patients without suicidal ideation (with score of 0 in all items of suicide), and depressed patients with suicidal ideation ($F_2=1.63$, $p= 0.204$). It is worth noticing that the sample size was not powered for the last analysis (n= 20, 18 and 23 respectively) (Cohen, 1992).

	Suicide risk			
Variable	Model 1		Model 2	
	B	β	B	β
Constant	-0.354**		0.234	
BDI Self-Esteem	0.395**	0.518	0.351**	0.461
D score			-0.306*	-0.238
R^2	0.269		0.322	
F	21.661**		13.753**	
ΔR^2	0.269		0.053*	
ΔF	21.661**		4.543**	

*p<0.05; **p<0.01

Table 23: Hierarchical multiple regression predicting suicidal risk from subscales of self-reported self-esteem, and implicit self-esteem

Discussion

This chapter attempted at contributing to the knowledge about the relevance and usefulness of the SE-IAT to differentiate between depressed and not depressed patients, and to predict depressive features in a unique way. As it was reviewed in the introduction to this chapter, the relationship between implicit self-esteem and depression is contradictory among studies, so replications are warranted. The present study also sought to determine the direction of the contribution of implicit self-esteem to depression, considering the abundance of contradictory results in that regard.

Implicit self-esteem as measured with the IAT failed to discriminate between depressed patients and controls, therefore this study corresponded to the third group mentioned in the introduction (with others like Franck et al., 2008). In the full sample, implicit self-esteem was positive. That is to say, people responded faster to congruent trials than to those that were incongruent. Most papers that describe implicit self-esteem as measured with the IAT and report *D* scores, find this positive trend (De Jong et al., 2012; De Raedt et al., 2006; Franck et al., 2008; Steinberg, 2006; Steinberg et al., 2007; Wegener et al., 2015). A study exploring this phenomenon used the SE-IAT, but the "other" category was not an undifferentiated other, but someone from the in-group of

the participant, and his or her best friend in three samples: an American, a Japanese, and a Chinese sample. Positive self-esteem scores were found throughout the samples, which suggest that implicit self-esteem is universally positive (Yamaguchi et al., 2007). This idea is ratified by the studies of Remue and colleagues, who replicated the null finding of a similar level of positive implicit self-esteem in both dysphoric and non-dysphoric patients. They found that these two groups could only be discriminated by the difference between their actual self-esteem ("I am valuable") and their ideal self-esteem ("I want to be valuable"). Dysphoric patients showed a significantly larger discrepancy between actual and ideal self-esteem (Remue, De Houwer, Barnes-Holmes, Vanderhasselt, & De Raedt, 2013; Remue et al., 2014). As we have seen in the previous chapter, self-esteem can be understood as a comparative attitude, where one's own value is always in reference to a real or imagined other (Greenwald & Banaji, 1995).

In spite of the present study's null finding regarding the difference between depressed participants and controls, implicit self-esteem was related to severity of depression. Higher implicit self-esteem reduces the odds of presenting with a more severe degree of depression, according to the HDRS.

The relationship between implicit self-esteem and depression was inverse in most symptomatic measures. However, most of these correlations failed to reach significance, when this latter one was corrected using Bonferroni adjustments. It is to be expected that bigger samples will find more statistically significant inverse relationships between self-esteem and depression, indicating that implicit self-esteem works similarly to its explicit counterpart, with depressive individuals showing a trend to present lower implicit and explicit self-esteem. It is a known finding that explicit and implicit self-esteem are more related to each other when the focus is affective in comparison to a cognitive thematic focus (Smith & Nosek, 2011). The setting of this experiment implied an affective focus, whereby people were asked about their feelings towards the world, the future and the self. While explicit attitudes could be separated into domains that are either more cognitive or affective, implicit attitudes are more automatic, that is to say, emerging more from affective states rather than from a rational processing of information. The relationship between negative implicit self-esteem and depression is further reinforced by the fact that when people tend to ascribe validity to their intuitive beliefs, the correspondence between explicit and implicit attitudes is increased (Jordan, Whitfield, & Zeigler-Hill, 2007). This implies that implicit self-attitudes are experienced as intuitive valuations of the self that are incorporated into explicit self-valuations.

What is more interesting is the relationship found between implicit self-esteem and self-reported suicidal ideation. Even when controlling for proxies of explicit self-esteem, implicit self-esteem was able to predict suicidal ideation in the same direction as explicit self-esteem, namely lower self-esteem predicts higher levels of self-reported suicidal ideation. It is noticeable that implicit and explicit self-esteem were not related to each other. This finding must be interpreted with care. Suicidal ideation predicted by self-esteem was obtained through self-report, which is in itself only partially useful for the prediction of actual suicidal behaviour (Busch et al., 2003). Also, this study failed to replicate (and did not have enough statistical power to reliably do so) a discriminating difference between suicidal and non-suicidal depressive patients as Franck and colleagues were able to do (Franck et al., 2007a). In fact, the predictive power of lower self-esteem runs against the findings of Franck and colleagues, who found that implicit self-esteem was significantly higher in depressed patients with suicidal ideation than those without, to the extent that there were no significant differences in implicit self-esteem between depressed patients with suicidal ideation and non-depressed controls. An exploration of the relationship between self-esteem and suicide with a bigger sample is warranted to help discard the possibility that this might have been a chance finding due to the several analyses carried out on this data.

Limitations

When addressing the limitations of this study, we need to keep in mind its exploratory nature. In this sense, the most important limitation of this study is the small sample size, which affected the statistical power of the study to carry out analyses of interest. For example, the replication of studies analysing group differences between depressive patients with and without suicidal ideation (Franck et al., 2007a). Along the same lines, in connection to suicide and depression, a bigger sample could relate symptomatic measures to the discrepancy between implicit and explicit self-esteem and the direction of such discrepancy. Underpowered analyses on this sample (not reported) showed promising and interesting results in relation to the implicit-explicit discrepancy, which is a research domain that could also benefit from replication (Creemers, 2014; Creemers, Scholte, Engels, Prinstein, & Wiers, 2012; Creemers et al., 2013; De Jong et al., 2012; Dentale, Vecchione, Coro, & Barbaranelli, 2012; Dimaro et al., 2015; Franck et al., 2007a; Jordan et al., 2007; Nakamura et al., 2015; Olson et al., 2007; Schreiber, Bohn, Aderka, Stangier, & Steil, 2012; Valiente et al., 2011; Zeigler-Hill, 2006).

Another important limitation in the findings of this study is the self-reported nature of most measures (with the obvious exception of the IAT), particularly those related to suicide. This implies that the interesting findings regarding the prediction of depression by both implicit an explicit self-esteem must be treated with suspicion, in light of the fact that most people who engage in suicidal behaviour did not express their intentions explicitly, and even deny them when prompted directly (Busch et al., 2003).

Finally, the battery of measures did not include measures of explicit self-esteem, thus a proxy variable had to be created from the BDI-II. In spite of the fact that the BDI-II shows very high correlations with specialised measures of explicit self-esteem, the use of only four items to construct that variable will inevitably miss many of the nuances and richness of the self-esteem construct that a specialised measure of self-esteem would provide.

In conclusion, this study found that implicit self-esteem is different from explicit self-esteem, although in depression it functions in the same direction as explicit self-esteem, namely lower explicit and implicit self-esteem seem to be markers of depression. However, implicit self-esteem is not sufficient to discriminate between depressed and non-depressed patients, but it is related to severity of depression and suicidal ideation. The next chapter will attempt to replicate these findings with a larger sample and explore the relationship between implicit/explicit self-esteem discrepancy and symptomatic measures of depression.

Chapter 5: Implicit Self-Esteem in the Context of Intervention. Baseline Analysis

Introduction

The empirical study reported in the previous chapter found only limited support for the notion of implicit self-esteem as a correlate of depression. Implicit self-esteem did not differentiate between depressed patients and healthy controls. In the full sample, implicit self-esteem was related to severity of depression and suicide. The present chapter reports a study on a larger sample composed only of depressed patients, with the objective of following up and replicating the previous findings. Also, to add a new variable: self-esteem discrepancy, namely the difference between explicit and implicit levels of self-esteem, and its relationship to symptomatic outcomes in depression. Although scarce, there are interesting results in the study of self-esteem discrepancy.

Despite the fact that it is commonly known that high explicit self-esteem is related to positive outcomes (Tennen & Herzberger, 1987), it has also been found to be related to negative behaviours such as prejudice, aggression, and higher levels of defensiveness (Baumeister et al., 2003; Baumeister, Smart, & Boden, 1996; Crocker, Thompson, McGraw, & Ingerman, 1987; Jordan et al., 2003; Kernis, 2003; Sandstrom & Jordan, 2008; Verkuyten, 1996; Verkuyten & Masson, 1995). Theorists have solved this apparent contradiction by postulating the existence of two different forms of high self-esteem: secure and fragile. Secure self-esteem is high but realistic, stable and robust to daily threats, while fragile high self-esteem reflects high but vulnerable feelings of self-worth that are susceptible to daily challenges, a need for constant validation and even a degree of self-deception (Rogers in Zeigler-Hill, 2006).

One of the ways to distinguish secure from defensive self-esteem is to compare explicit and implicit self-esteem. A person who presents a high level of self-reported self-esteem, but low levels of implicit self-esteem, is said to have a fragile self-esteem (Bosson et al., 2008; Campbell, Bosson, Goheen, Lakey, & Kernis, 2007; Vater et al., 2013). Fragile high self-esteem is associated to narcissism (Bosson et al., 2000), self-esteem instability and inappropriate levels of self-enhancement and defence in every day threats to the self (Bosson et al., 2008; Gregg & Sedikides, 2010; Schröder- Abé et al., 2007b; Vater, Schröder-Abé, Schütz, Lammers, & Roepke, 2010; Zeigler-Hill, 2006). For example, Bosson and her team reported two studies which found that people with high explicit self-esteem and low implicit self-esteem (measured with the Name Letter Task) display more unrealistic optimism, they identify more with an

unrealistically positive personality profile, and show reduced discrepancy between the ideal self and the actual self. The authors conclude that people with fragile self-esteem are more prone to higher levels of self-enhancement, compared to those with both high implicit and explicit self-esteem (Bosson, Brown, Zeigler-Hill, & Swann, 2003). People with discrepant high self-esteem have insecurities and self-doubts at a non-conscious level, which might lead them to bolster their feelings of frailty with explicit grandiosity. Fragile self-esteem is also unstable and context-dependent, which explains the nned for constant self-enhancement and the notion that, for example, narcissistic patients are fragile and volatile (Brown & Bosson, 2001; Gregg & Sedikides, 2010). In a sample of children aged 12-14 years, it has been reported that a pattern of high explicit self-esteem and low implicit self-esteem is positively associated with physical and relational aggression in the classroom, as reported by teachers (Sandstrom & Jordan, 2008). Besides externalising problems, a pattern of fragile self-esteem has also been reported in paranoid patients (Valiente et al., 2011).

However, it would be mistaken to understand the flux of self-esteem levels solely as negative instability. Implicit self-esteem can be flexible in an adaptive manner. People with high levels of explicit self-esteem, when their self-concept is threatened, display an increase in implicit self-esteem, as measured by the Name Letter Preference Test in experimental settings, while people with low levels of explicit self-esteem display a decrease in implicit self-esteem under the same threats (Jones et al., 2002). This same pattern was observed in a longitudinal naturalistic setting, where the threat to the self was negative events from normal day-to-day life (DeHart & Pelham, 2007). In this sense, it has been reported that people who are action-oriented show more independence of their implicit self-esteem levels from day to day challenges (Koole, 2004). Action orientation refers to the capacity of an individual to activate their extension memory, "a central executive system that supports intelligent, high-inferential form of intuition, and is characterised by the activation of extended networks of implicit self-representations" (p. 103). People who are capable of resorting to internal/implicit self-views when facing a threat to the self show more stable levels of implicit self-esteem thus avoiding potentially undermining influences of external challenges. This is radically different from the reinforcement of explicit self-esteem in narcissistic people who, when threatened, tend to show explicit grandiosity (Bosson et al., 2008; Gregg & Sedikides, 2010; Vater et al., 2013; Zeigler-Hill, 2006). Taken together, these findings seem to suggest that the levels of implicit self-esteem are flexible for some people (high explicit self-esteem or action-oriented people), and play a role in stress regulation.

The other configuration of self-esteem, known as damaged self-esteem, is characterised by a low level of explicit self-esteem and concurrent higher levels of implicit self-esteem. This kind of discrepancy has been less studied. Discrepant self-esteem, both fragile and damaged, has been related to increased anger suppression (feeling but not showing anger), more depressive attributional style, more nervousness, more self-doubt and more days of impaired health (Briñol, Petty, & Wheeler, 2006; Schröder- Abé, Rudolph, & Schütz, 2007a). In particular, damaged self-esteem has been related to higher levels of alexithymia (Dentale, San Martini, De Coro, & Di Pomponio, 2010), suicidal ideation, depressive mood and loneliness (Creemers, 2014; Creemers et al., 2012; Creemers et al., 2013; Kesting et al., 2011; Pavlickova et al., 2014). However, as reviewed above, high implicit/low explicit self-esteem also has a defensive function: healthy people with damaged self-esteem tend to interpret ambiguous statements as more positive than people with both low explicit and implicit self-esteem, and seem to pay more attention to positive than negative feedback (Schröder-Abé et al., 2007b).

The discrepancy between explicit and implicit self-esteem appears to be moderated by dismissive attachment. A study found that indices of dismissive attachment positively correlated with self-esteem discrepancy, for implicit self-esteem higher than its explicit counterpart and marginally, for higher explicit than implicit self-esteem (Dentale et al., 2012). This makes sense in the light of attachment literature, which posits a difficulty with emotional awareness among dismissively attached people (Picardi, Toni, & Caroppo, 2005; Waller & Scheidt, 2006). In fact, negative emotion is heightened in people who present damaged self-esteem when they are made aware of themselves. An experimental design asked people with damaged self-esteem to complete a brief task while a mirror was left in the room. This manipulation increased negative affect only for those participants who had damaged self-esteem. When the mirror was not present, negative affect was kept constant in all participants (Cheng, Govorun, & Chartrand, 2012).

On the same line, damaged self-esteem is present in patients with social anxiety disorder, both in adults and adolescents (de Jong, 2002; De Jong et al., 2001; Schreiber et al., 2012; Tanner et al., 2006). In the adult studies, the level of implicit self-esteem was slightly lower in socially anxious patients than in controls, but their explicit self-esteem was considerably lower. The mechanism is illustrated in the adolescent study (Schreiber et al.), where participants were measured after a social threat (they were told they had to give a speech about themselves in front of a camera, to be evaluated by

experienced psychologists). After the threat, socially anxious adolescents presented even lower explicit self-esteem, while their implicit self-esteem remained high. Both levels of self-esteem remained high in controls.

In borderline patients, high implicit/low explicit self-esteem is associated with severity of borderline symptomatology, while fragile and congruent self-esteem is not. Both fragile and damaged self-esteem patterns predicted heightened levels of self-harm in these patients. Dysphoria was predicted only by fragile self-esteem, and was unrelated to damaged self-esteem (Vater et al., 2010). It seems that the damaged self-esteem pattern, with a heightened automatic self-evaluation and a diminished reflexive self-evaluation is more akin to impulsivity than to low mood. The same pattern of damaged self-esteem has been found in a clinical sample of people with bulimia nervosa or binge eating disorder, known for their perfectionism and impulsiveness (Cockerham et al., 2009).

Regarding suicide, a comparison between depressive patients with suicidal ideation, depressive patients without suicidal ideation and not depressed controls showed interesting results. While controls showed normal levels of both implicit and explicit self-esteem, depressed patients without suicidal ideation showed both low explicit and implicit self-esteem. Depressed patients with suicidal ideation showed low explicit self-esteem, similar to depressed patients without suicidal ideation, but increased implicit self-esteem, similar to healthy controls. Implicit self-esteem was not related to other symptoms of depression. The discrepancy between implicit and explicit self-esteem was specific to suicidal ideation (Franck et al., 2007a). A similar pattern of associations was reported in two young females samples, one using the Name Letter Preference and the other using the SE-IAT (Creemers et al., 2012; Creemers et al., 2013). While implicit self-esteem was not related to depressive symptoms, damaged self-esteem discrepancy was a predictor of suicide ideation, depression and loneliness. Given that higher implicit rather than explicit self-esteem is also an indicator of self-esteem instability (Bosson et al., 2000), this discrepancy might be reflecting instable self-esteem, which is also associated with suicide as a moderator of the level of explicit self-esteem (De Man & Gutierrez, 2002).

The reader must have already noticed that the literature on fragile self-esteem seems to convey that fragile self-esteem fosters defensive reactions and explicit self-enhancement as a way of avoiding a deeper and more automatic feeling of inadequacy. While on the other hand, the literature on damaged self-esteem seems to convey the

presence of negative affect stemming from a sense of incongruence between automatic grandiose self-attitudes and a more real, but lower explicit self-esteem. Both types of discrepancies have the potential to be maladaptive, but in the case of damaged self-esteem, it has been theorised that this discrepancy might be representing the incongruence between an ideal self (implicit self-esteem) and the actual self (explicit self-esteem). People with damaged self-esteem feel trapped between goals and reality, which is experienced as disturbing (Franck et al., 2007a).

It has been reported that the relationship between explicit and implicit self-esteem is moderated by gender. Women would tend to show more congruent self-esteem, that is to say more similar levels of implicit and explicit self-esteem than men. The authors explain this as a reflection of the fact that women are more socialised to trust their intuition than men. This finding must be interpreted with care, given that a substantial difference was found between men and women when conflating the results of six different samples, but most of these individual samples did not achieve significance in the gendered relationship between implicit and explicit self-esteem (Pelham et al., 2005). However, it has been found that people who see their intuition as valid (regardless of gender) show more congruent levels of implicit and explicit self-esteem (Jordan et al., 2007).

Hypotheses

Following this brief summary of the existent literature, and taking into account the findings of the previous chapter, it is expected that implicit self-esteem, measured with SE-IAT D scores will not be related to general measures of depression. However, they will have an incremental predictive power over explicit measures of self-esteem in the severity of depression.

It is expected that D scores will correlate with measures of suicidality. However, the direction of this correlation is difficult to surmise. While in the literature high levels of implicit self-esteem are related to more severe suicidal ideation, in the last chapter of this thesis the opposite result was obtained.

Regarding self-esteem discrepancy, it is expected that women will show a more congruent self-esteem configuration than men. It is also anticipated that larger values of self-esteem discrepancy will predict depressive mood and suicidal ideation. The expected direction of the discrepancy is towards damaged self-esteem, namely higher implicit and lower explicit self-esteem.

Methods

Design

This is a cross-sectional correlational study. In order to test the hypotheses, the depressed adult participants of the previous study (fMRI) were merged into the baseline of the REDIT study (details below and in Chapter 6). The merged sample had completed the same SE-IAT together with other validated measures tapping specific symptomatic components of depression.

Procedure

For the procedure followed by the fMRI subsample, please see Chapter 4. The rest of the depressed participants were part of an outcome study, the Randomised Evaluation of Dynamic Interpersonal Therapy (REDIT). Dynamic Interpersonal Therapy is a weekly individual psychodynamic psychotherapy for adult depression with a total duration of sixteen sessions, based on theoretical notions of attachment and mentalizing (Lemma, Target, & Fonagy, 2011), which was contrasted to an enhanced waiting list (a 16-week waiting list with fortnightly low-intensity controls by a health professional). Participants were identified at triage/initial assessment at the two participants sites in London. Participants were then informed of the trial and were given the choice of being contacted by a research assistant. If the participant consented, they were contacted after 48 hours, to arrange an assessment meeting. Participation in the study was completely voluntary and informed consent was given in written. Participants completed baseline measures before randomisation, with research assistants trained in the administration of various measures. The data to be analysed in this chapter corresponds to the baseline data of all participants included in the study, regardless of subsequent randomisation. The next chapter will deal with follow-up measures, taken at mid-therapy and at the end of treatment (8 and 16 weeks, respectively). Patients were paid for each of the data collection sessions.

Participants

All participants needed to be 18 years or older and have a diagnosis of depression (with or without dysthymic disorder), confirmed by a score of 14 or above in the Hamilton Depression Rating Scale (HDRS). REDIT participants needed a confirmation

by triage of their need for high-intensity treatment. Participants needed to be fluent in English.

Participants were excluded if they presented current bipolar disorder or psychotic symptoms, current use of antipsychotics, severe personality disorder, eating disorder, historic or current self-injury/parasuicide or current excessive use of drugs or alcohol. REDIT participants were excluded if they had participated in another depression clinical trial in the previous year where the participant had received cognitive-behavioural therapy (CBT), or a history of unsuccessful CBT treatment. They were excluded from REDIT if they presented a clinical contra-indication to short term psychotherapy (e.g attachment history with multiple separations, serious prolonged trauma in childhood, multiple caregivers), and highly or unstable life arrangements (such as homelessness).

Exclusion criteria from the fMRI subsample also took into account counter indications to magnetic resonance scanner (see previous chapter).

The final sample was composed of 147 depressed participants, 107 from the REDIT trial and 40 from the fMRI sample. One participant was concurrently taking part of the REDIT study and the fMRI study. His measures were kept in the REDIT subsample.

Measures

The REDIT subsample completed other measures besides the following ones. However, they will be described in detail in the next chapter, given that they are irrelevant for the fMRI subsample.

Implicit Self-Esteem

Implicit self-esteem was measured with the Self-Esteem Implicit Association Test (SE-IAT; Greenwald et al., 1998) adapted for iPad. This measure has been described in detail in previous chapters. The SE-IAT yields a final score known as *D*, with larger values indicating higher levels of implicit self-esteem.

Depression

Two measures were used to measure depressive symptoms. The Beck Depression Inventory, Second Edition (BDI-II; Beck et al., 1996), and the Hamilton Depression Rating Scale (HDRS; Hamilton, 1960). Both have been described in the previous chapter. They are both largely used in research and clinical settings, with adequate to excellent psychometric indicators.

General Psychopathology

To evaluate this sample's general psychological distress and psychiatric disorders the Brief Symptom Inventory (BSI; Derogatis, 1993) was used. This is a 53-item, 5-point Likert self-report scale ranging from "not at all" to "extremely", which takes approximately 4 minutes to administer. It is based on a longer measure, the Symptom Check-List-90-R, and it is considered by its authors to be the short version of that measure (Derogatis, 1992). It comprises nine primary symptom dimensions or subscales: somatisation (reflects psychological distress arising from perception of bodily dysfunction), obsessive-compulsive (thoughts and actions experienced as unremitting and irresistible, but are ego-dystonic), interpersonal sensitivity (feelings of personal inadequacy and inferiority, and discomfort during interpersonal interactions), depression (symptoms of dysphoric affect and mood, and loss of vital energy), anxiety (restlessness, nervousness, tension and panic), hostility (thoughts, feelings and actions of aggression and destruction), phobic anxiety (fears oriented towards travel, open spaces, crowds, public places, etc.), paranoid ideation (conceived as a mode of thinking which uses projection, hostility, suspiciousness, centrality and fear of loss of autonomy) and psychoticism (from signs of schizoid, alienated lifestyle to more dramatic symptoms of psychosis). Besides the subscale score, the BSI yields three global indices of general psychopathological distress. This measure is widely used in research, given its brevity and excellent psychometric properties. The instrument has shown good internal reliability scores averaging Cronbach's $\alpha = 0.75$ and tests-retest reliability averaging $r = 0.8$ (Derogatis & Melisaratos, 1983). This measure was only completed by the REDIT subsample.

Data Analysis

Data was analysed using SPSS 21 and 22 for Macintosh computers. After descriptive exploration of demographic and symptomatic variables, transformations will be carried out to achieve normal distribution of variables when needed. Variables that do not achieve normality after transformations will be analysed using non-parametric tests. Correlations will be used to explore the relationships between variables. Significant correlations will be followed-up by regression analyses. ANOVA and t-tests will be utilised to explore group differences in D scores for categorical variables.

New variables will be created. A composite variable of suicide will be developed by standardising and averaging subscales for suicide from the HDRS and the BDI-II. A

dichotomous variable indicating presence or absence of suicidal ideation in the HDRS will also be calculated. A proxy of explicit self-esteem will be created from the self-esteem items of the BDI-II, by reversing, standardising and averaging those items. *D* scores will be standardised and subtracted from the standardised proxy of explicit self-esteem. The resultant variable will represent the discrepancy between implicit and explicit self-esteem scores, with larger scores reflecting higher scores in explicit over implicit self-esteem. Correlations between this discrepancy score and other measures will be calculated, and significant relations will be followed-up with regression models.

Results

Preliminary Analyses and Demographics

The total final sample was composed of 147 participants, as shown in Table 24. The total scores for the BDI-II showed normal distributions. *D* scores were calculated using the same procedure described in Chapter 3 (Greenwald et al., 2003). *D* scores were moderately negatively skewed, so a reflected square root transformation was carried out, followed by an inversion of the transformed variable so that larger scores reflect higher implicit self-esteem (*Shapiro-Wilk* = 0.99, *p* = 0.348). The same transformation achieved normality for the BSI total score and the BSI-General Severity Index (*Shapiro-Wilk* = 0.98, *p* = 0.358 and *Shapiro-Wilk* = 0.98, *p* = 0.0350, respectively). Age was moderately and positively skewed, so a square root transformation was carried out. This transformation achieved normality for age

	Female (n = 97)	Male (n = 50)	Total (n = 147)
Age (SD)	36.75 (10.71)	37.53 (12.14)	37.02 (11.18)
HDRS (SD)	19.05 (4.70)	17.34 (2.80)	18.47 (4.21)
BDI-II	34.80 (8.17)	31.81 (8.59)	33.74 (8.41)
BSI	90.27 (37.66)	76.33 (28.03)	85.45 (35.08)
D score	0.094 (0.70)	0.347 (0.72)	0.239 (0.66)

Table 24: Means and standard deviations for age and measures by gender. HDRS = Hamilton Depression Rating Scale; BDI-II = Beck Depression Inventory, 2ⁿᵈ Edition; BSI = Brief Symptom Inventory; D score: Untransformed Self-Esteem Implicit Associations Test scores.

(*Shapiro-Wilk* = 0.99, *p* = 0.098). The mean *D* score was positive and significantly different from zero (t_{96} = 26.78, *p* < 0.001). *D* scores showed no association with age (*r* = -0.78, *p* = 0.45). The Spearman-Brown coefficient for split-half reliability on difference scores of odd trials against even trials for this SE-IAT was 0.880. Total scores for the HDRS were not normal and positively skewed. Nonparametric tests will be used with this variable. Genders showed no statistical differences regarding age (*t* = -0.405, *p* = 0.69). There were no differences for gender in the measures' total scores, including *D* score (*t* = 1.725, *p* = 0.088), except for the HDRS. For this last measure, a Mann-Whitney U test was run to determine if there were differences in the depression score between males and females. Distributions of the engagement scores for males and females were similar, as assessed by visual inspection. The Median engagement score was statistically significantly different between males and females, *U* = 1.943, *z* = -1.98, *p* = 0.048, using an exact sampling distribution for *U*. Depression scores on the HDRS were higher for women.

Symptomatic measures

Total scores for measures of psychopathology were significantly and positively correlated, however none of them is particularly strong, as shown in Table 25. It is noteworthy that the strong correlations between the BDI-II and HDRS normally found in the literature (e.g. Mandić-Gajić, Samardžić, & Špirić, 2015) failed to emerge in this sample, given the artificial truncation of the HDRS at 14 points, which was the main inclusion criterion for the REDIT study.

The total sample completed the HDRS. The distribution of depression severity according to the HDRS and the BDI-II is shown in Figure 13. Note that the BDI-II was

	1	2	3	4
1. BDI-II[a]	1			
2. BSI[a]	0.562**	1		
3. HDRS[b]	0.256**	0.484**	1	
4. D score	0.164[a]	-0.045[a]	-0.026[b]	1

**. Correlation is significant at the 0.01 level (2-tailed).
[a] = Pearson correlation. [b] = Spearman correlation

Table 25: Correlation coefficients between symptomatic measures

Figure 13: Severity of depression according to the HDRS-17 and the BDI-II
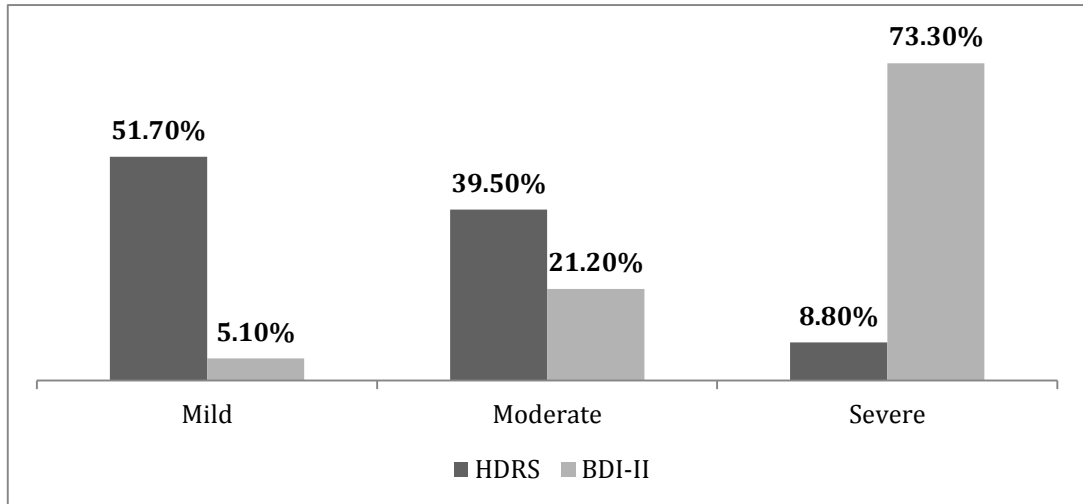
completed by 80.3% of the sample, and that the cut-off points are slightly different than the ones recommended for the HDRS. Separated analyses will be carried out for each depression measure, because of the medium correlation found between the two.

The SE-IAT was not completed by the full sample, but by 96 participants. Participants who did not complete the SE-IAT were all from the REDIT subsample. Comparisons between participants who completed the SE-IAT against those who did not, showed no significant differences in symptomatic scores, as assessed with t-tests.

As shown in Table 25, D scores were not related to total scores of the symptomatic measures, in line with the hypotheses. This lack of correlation appears also when relating D scores to subscales of the BDI-II, BSI and the HDRS. In particular, there was no relationship between D scores and BDI-II scores of self-esteem. Those relationships were also lacking in the previous chapter; and here again they were not found for these subscales, namely feeling like a failure (*rho* = 0.166, p = 0.161), self-dislike (*rho* = 0.017, p = 0.885), self-criticalness (*rho* = 0.050, p = 0.672), and worthlessness (*rho* = 0.156, p = 0.186).

More detailed analyses were performed. A One Way ANOVA was run to assess differences in D scores according to severity of depression as measured with the BDI-II. This test found no differences between severity groups in D scores $F_{2,72}$ = 1.002, p = 0.373. Regarding the HDRS, no relationships were found. A One Way ANOVA found no differences in D scores according to HDRS depression severity, $F_{3,93}$ = 0.368, p =

0.776. This null result remained after combining the severe and very severe groups into one single group, $F_{2,93} = 0.277$, $p = 0.759$.

D scores did not correlate with suicide scales ($r = 0.142$, $p = 0.376$; $r = -0.029$, $p = 0.777$ for the BDI-II and the HDRS, respectively). A composite suicide score was created by averaging the standardised scores of the suicide subscales of the HDRS and BDI-II ($\alpha = 7.81$), to which D scores were not related ($r = -0.006$, $p = 0.954$). As in the previous chapter, a hierarchical regression was carried out to determine if the addition of implicit self-esteem improved the prediction of self-reported suicidal risk over and above self-reported self-esteem alone. Standardised scores of the subscales of the BDI-II thought to reflect explicit self-esteem, namely feeling like a failure, self-dislike, self-criticalness, and self-worthlessness, were reversed and averaged, so higher scores will reflect higher explicit self-esteem. They showed an acceptable level of internal consistency (Cronbach's $\alpha = 0.762$) and were not related to D scores ($r = -0.107$, $p = 0.368$). This proxy of explicit self-esteem was inversely correlated with suicide ($r = -0.343$, $p < 0.001$). The linear regression model including both explicit and implicit self-esteem was statistically significant ($R^2 = 0.459$, $F_{(2, 70)} = 9.364$, $p < 0.001$; adjusted $R^2 = 0.189$). However, the addition of implicit self-esteem to the prediction of suicide risk did not achieve a statistically significant increase in R^2 (0.053, $F_{(1, 70)} = 4.543$, $p = 0.636$).

Explicit/Implicit Self-Esteem Discrepancies

To calculate an index for the discrepancies between explicit and implicit self-esteem, standardised scores for D were also calculated. Then, the standardised D score was subtracted from the averaged standardised score for explicit self-esteem, so positive values represent higher explicit than implicit self-esteem. Self-esteem discrepancies can be increased in both directions: if it is increased towards a higher absolute value with a positive sign, then explicit self-esteem is higher than implicit self-esteem and vice versa, when the discrepancy value is increased towards a higher absolute value with a negative sign, then implicit self-esteem is higher than implicit self-esteem.

See Figure 14: the average self-esteem discrepancy for people who mention suicide only when questioned is smaller than that of participants who report suicide ideation spontaneously as a severe symptom. Discrepancy between explicit and implicit self-esteem is larger in the more severely ill participants, and its negative sign indicates us that the direction of the discrepancy is a comparatively higher implicit over explicit self-esteem. The direction of the discrepancy is told by the sign of the discrepancy value. This discrepancy score was not related to gender but, as hinted in the previous chapter, this score was significantly associated with several symptomatic scales. Increased discrepancies between explicit and implicit self-esteem, favouring implicit over explicit self-esteem were related to increased scores on the HDRS subscale of depressed mood ($rho = -0.299$, $p = 0.010$) and suicide ($rho = -0.347$, $p < 0.003$). The same pattern of association is found between self-esteem discrepancy and the depressive mood subscale of the BSI ($rho = -0.438$, $p < 0.004$).

Given these results, more minute analyses were carried out. Self-esteem

Figure 14: Means for explicit and explicit self-esteem, and their discrepancy, according to the depressed mood subscale of the HDRS
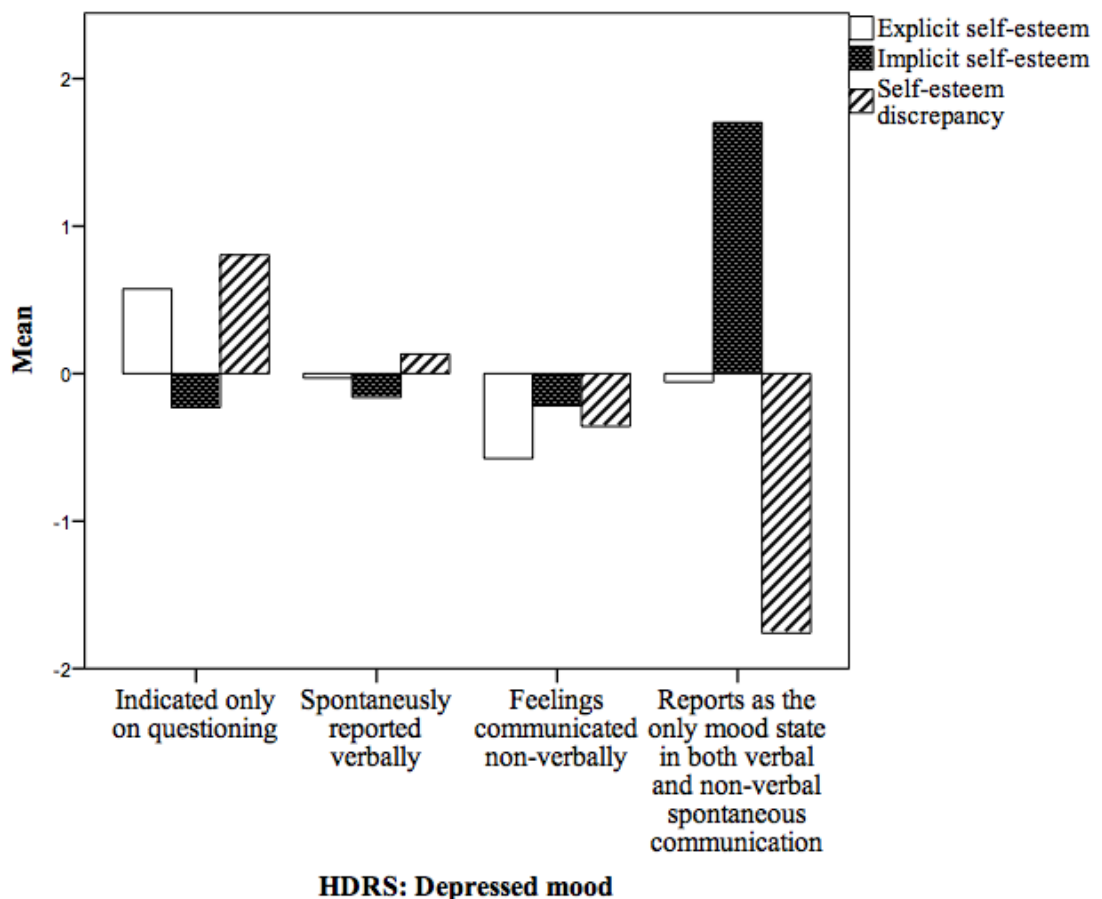
discrepancy was entered as a predictor of the suicide composite variable in a linear regression that was a good fit for the data, $F = 5.439$, $p = 0.006$, explaining 13.4% of the variance. Suicide levels increased when the discrepancy increased, with implicit self-esteem higher than explicit self-esteem, $\beta = -0.236$, $p = 0.003$. A significant linear prediction of the depressive subscale of the BSI by implicit/explicit self-esteem discrepancy was found $F(1,39) = 8.684$, $p = 0.005$. Self-esteem discrepancy accounted for 16.10% of the explained variability in depression, as measured by the BSI. A larger discrepancy, favouring implicit over explicit self-esteem was found to predict more depressed mood ($\beta_0 = 2.257$; $\beta_1 = -0.234$; all $p \leq 0{,}005$).

A cumulative odds ordinal logistic regression with proportional odds was carried out, to estimate the effect of self-esteem discrepancy on depressed mood as measured by the HDRS. The assumption of proportional odds was met, as assessed by a full likelihood ratio test comparing the residual of the fitted location model to a model with varying location parameters, $\chi^2_{(2)} = 1.015$, $p = 0.602$. This model statistically significantly predicted depressive mood as measured by the HDRS, over and above the intercept-only model $\chi^2_{(1)} = 6.726$, $p = 0.009$. This model found that an increase in the discrepancy between implicit and explicit self-esteem (with implicit self-esteem higher than explicit self-esteem) was associated with an increase in the odds of patients reporting feelings of sadness, and depression as virtually their only feelings, with an odds ratio of 0.644 (95% CI, 0.455 to 0.911), Wald $\chi^2_{(1)} = 56.355$, p = 0.013. Figure 14 shows the means for implicit and explicit self-esteem, as well as their discrepancy.

The same analysis was carried out for the suicide subscale of the HDRS. The assumption of proportional odds was met, $\chi^2_{(2)} = 4.681$, $p = 0.096$. This model was a good fit for the data and significantly predicted the odds of presenting suicidal ideas or gestures as measured by the HDRS, over and above an intercept-only model $\chi^2_{(1)} = 9.542$, $p = 0.002$. An increase in the discrepancy between explicit and implicit self-esteem (with implicit higher than explicit) predicted an increase in the odds of reporting suicidal ideation or gestures, with an odds ratio of 0.594 (95% CI, 0.421 to 0.838), Wald $\chi^2_{(1)} = 8.802$, p = 0.003. Figure 15 shows the means for explicit and implicit self-esteem, and their discrepancy, grouped according to the suicide subscale of the HDRS. Furthermore, patients with any level of suicidal ideation (with a score > 0 in the HDRS suicide subscale) scored significantly different in measures of self-esteem, as shown in Table 26.

Figure 15: Means for explicit and implicit self-esteem, and their discrepancy, according to the suicide subscale of the HDRS
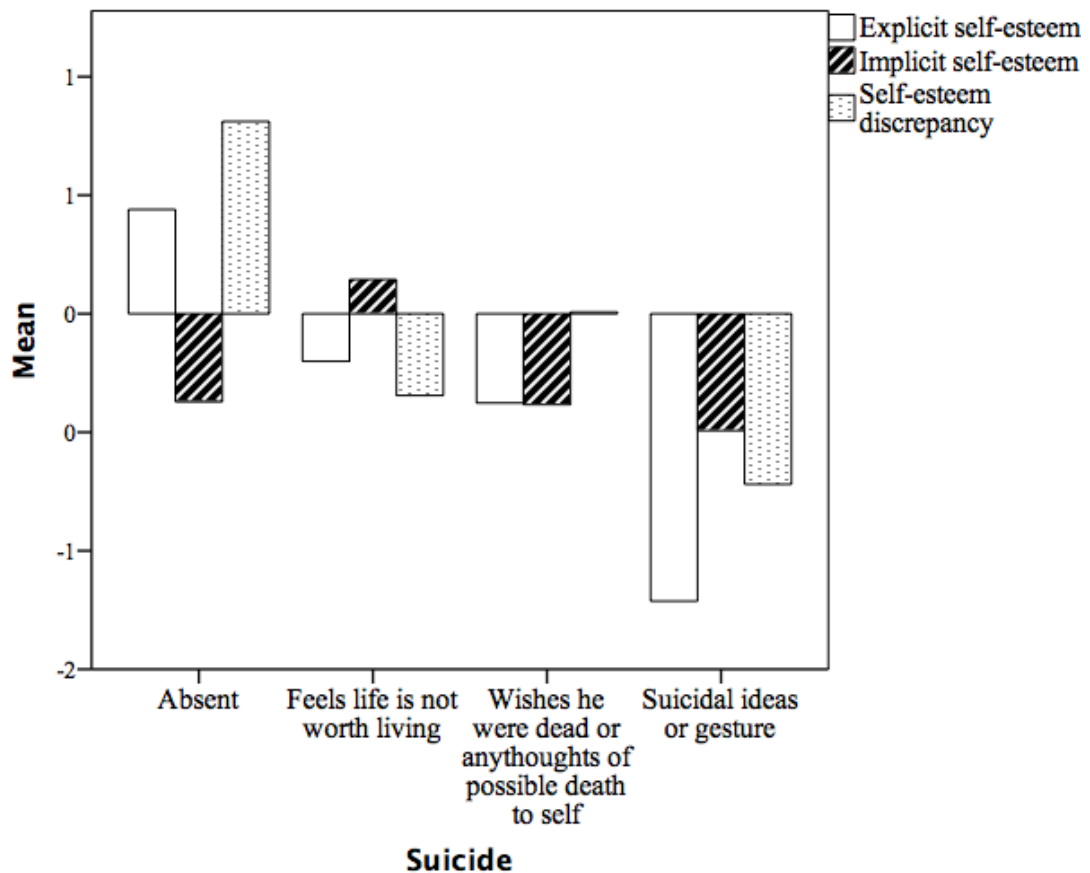
		Mean	Std. Deviation	*t*
Explicit self-esteem	no suicidal	0.2585	0.70053	2.887**
	suicidal	-0.1525	0.76796	
Implicit self-esteem	no suicidal	0.4218	0.23680	-2.574**
	suicidal	0.5683	0.21697	

** *p* < 0.01

Table 26: Means, standard deviations and mean differences between suicidal and non-suicidal groups in different measures of self-esteem.

Discussion

This study aimed at discovering associations between symptomatic measures of depression and implicit self-esteem. It also hypothesised an effect of self-esteem discrepancy, or the difference between explicit and implicit proxies of self-esteem, on measures of depression and suicide. We have reviewed in the introduction to this chapter that the literature about self-esteem discrepancy has more consistent findings than those studies relating depression to implicit self-esteem alone.

Once again, *D* scores of implicit self-esteem were unrelated to total scores for the symptomatic measures, and to suicide subscales, in contrast to the results in the previous chapter. It is possible that the finding of the last chapter was a chance finding, as it was suspected. However, interesting results appeared when addressing self-esteem discrepancy. The previous chapter showed an inverse relationship for self-esteem, both explicit and implicit, and self-reported suicidal thoughts. In this chapter, the relationship was only significant for explicit self-esteem. However, the analysis of the discrepancy between implicit and explicit self-esteem showed that a larger discrepancy between implicit an explicit self-esteem, favouring higher implicit than explicit self-esteem, with both scores indicating negative explicit and implicit self-esteem predicts the presence of suicidal thoughts and gestures. In that sense, and considering the incremental predictive value of implicit self-esteem to explicit self-esteem on suicide found in the previous chapter, it seems that for suicide, implicit and explicit self-esteem must be considered together, as it has previously been found in the literature (Franck et al., 2007a). The discrepancy between explicit and implicit self-esteem was not related to gender, in contrast to the studies by Pelham (2005).

A way to understand the relationship between low explicit self-esteem, high implicit self-esteem and suicide comes from the dual-process framework of goal pursuit and goal attainment (Brandtstädter & Rothermund, 2002; Rothermund & Brandstädter, 2003). After successive failures in attaining a goal, usually expectations adapt to reality (accommodative coping), or an active solution is found to change reality (assimilative coping). The first might be the case of the depressed non-suicidal subsample, in the sense that they might not have goals to strive for, depression ensuing. But the sub-group of depressive patients with suicidal ideation has been described as showing a combination of unusually high standards and recent failures (Baumeister, 1990). It is possible to speculate that such high standards might be related to implicit positive self-attitudes, while explicit self-esteem represents perceived reality. When people feel entrapped between unrealistically high expectations and reality, they might engage in suicidal ideation as a way to actively escape this incongruence by assimilation, or active coping.

This notion is supported by findings that damaged self-esteem is associated with maladaptive forms of perfectionism, characterised by high and rigid standards and goals (Zeigler-Hill & Terry, 2007). People with damaged self-esteem also present higher levels of depressive and negative rumination than people presenting other self-esteem configurations (Phillips & Hine, 2014). In this sense, damaged self-esteem can give way to psychopathology when a person has a negative experience, or receives negative feedback about themselves. People with damaged self-esteem tend to show less depressive symptoms and higher global self-esteem after receiving positive feedback, but fare worse than controls when negative feedback is received, with a decrease in explicit self-esteem and an increase in depressive feelings (Jordan et al., 2013), and rumination. If we think that damaged self-esteem is also connected to impulsive auto-aggression, it can be concluded that facing negative life events can be a trigger for suicide in people with high implicit and low explicit self-esteem.

Most of the findings extracted from the literature, and those obtained here have been replicated using different measures of implicit self-esteem, giving these measures validity and support. The literature on self-esteem discrepancies is more harmonious than the literature trying to predict depressive symptomatic outcomes from implicit self-esteem alone. In the previous chapter we saw that there are at least three groups of disparate results within this literature when attempting to use measures of implicit self-esteem to distinguish between depressed participants and controls. It has also been observed throughout this thesis that implicit self-esteem, as measured by the SE-IAT,

does not yield consistent results when used as a criterion variable (van Tuijl et al., 2014). The usefulness of the construct of implicit self-esteem and the instruments trying to measure it become valuable when they are used as a moderator variable in general, or specifically in relation to explicit self-esteem. The same can be said about explicit self-esteem, as discussed at the beginning of this chapter: high explicit self-esteem seems to be associated to both positive and negative outcomes, which appear inconsistent. However, when the effect of self-esteem is understood through a specific configuration of implicit and explicit self-esteem, results are more consistent and lend themselves to wider and richer interpretations.

Limitations

Given that this sample includes the baselines of the REDIT study, which had an inclusion criteria of HDRS > 14 that includes people presenting moderate to severe depression. Such truncation of the measure implied that certain expected relationships were not found or, when found, were weaker than expected, as occurred with the weak correlation with the BDI-II. However, this is expected when studying a sample consisting only of clinical subjects, and we see that, in spite of this fact, many of the expected results emerged.

The sample in this study was larger than in the previous one. However, there were quite a few missing SE-IATs, mostly due to some participants taking too long in completing the HDRS interview, which is the primary outcome of the REDIT study and as such, it was collected first. This could imply that patients who did not complete the SE-IAT were more severely depressed, with higher levels of psychomotor retardation. However, participants who did not complete the SE-IAT scored similarly in symptomatic measures to participants who completed the implicit self-esteem measure.

When interpreting the results of this study, it is important to keep in mind that discrepancies between explicit and implicit measures of self-esteem might be reflecting a self-underpresentation in the explicit measures. A study showed that people with high explicit self-esteem and low implicit admitted over-presenting themselves at the end of the study. On the same line, when people are urged to avoid over- or under-presenting themselves in the explicit measures, the discrepancy between implicit and explicit tends to fade (Olson et al., 2007). People can indeed manipulate their explicit presentation if they are motivated to do so. In the case of this study, the REDIT subsample might have been motivated to present themselves as more depressed, in order to receive prompt

treatment. In the particular case of this study, however, people were informed of the randomisation process prior to completing baseline measures, thus aiding to dispel the need to present oneself as more severely depressed.

Chapter 6: Implicit Self-Esteem in the Context of Intervention. Longitudinal Analysis

Introduction

Throughout this thesis, we have been revising the pertinence of the iPad version of the Self-Esteem Implicit Association Test (SE-IAT) to different interpersonal domains of psychological research, specifically to depression. It is the objective of this last empirical chapter to address the response of this measure to the passage of time and the influence of psychotherapeutic intervention in a sample of depressed patients within the context of a randomised controlled trial (RCT).

Regarding the temporal stability of the Implicit Association Test, understood as test-retest reliability, reports indicate a stable nature of scores. Reliability indices of the measure are informed as the correlation between test results over a definite span of time.

These correlations differ in the literature, and they seem to be a function of the underlying construct being measured, as well as experimental manipulations and the presence of recent (negative) life events. Even considering that temporal stability of the IAT can vary according to content (and that such variation is still unknown), the analysis of several studies with different IATs and different time spans between administrations showed what authors consider a generally stable test-retest reliability index (median $r = 0.56$). Such general pattern is barely affected by the time span between administrations ($r = -0.0003*day + 0.568$) (Nosek et al., 2007a). Authors that have developed the test, or used it from early in its history, tend to judge this index as a reliable one. However, in general psychometrics, this index is considered poor (Bosson et al., 2000).

In spite of this general result, there are differences between studies that are not straightforward, which precludes from giving an overall reliability index. In the case of implicit prejudice, for example, for an IAT measuring attitudes towards homosexual people, the found correlation between two counterbalanced administrations the first day was $r = 0.55$. There was no experimental manipulation between administrations (Banse et al., 2001). On the other hand, a black/white racial IAT, administered twice in a 24 hour span, showed a correlation of $r = 0.65$, in spite of experimental manipulation. The experimental manipulation consisted in showing photos of admired black people after the first administration of the racial IAT. Such manipulation elicited significant differences in the implicit preference of white versus black people. Nevertheless the

IAT remained remarkably stable across the two administrations (Dasgupta & Greenwald, 2001). A similar study, utilising an IAT measuring self-association with anxiety, showed that an experimental manipulation (announcement to present a speech in public) did not affect IAT scores in comparison with a control group, while self-reported anxiety increased. A second study reported in the same paper, found a slight increase in implicit anxiety in the experimental condition, and a decrease in rest-retest reliability indices between the two administration, but only for the experimental group, which shows that the IAT responds to intervention (Schmukle & Egloff, 2004).

Three studies (with three different samples) reported by Egloff, Schwerdtfeger and Schmuckle (2005) administered an Anxiety-IAT (associations of self with words related to anxiety vs calmness) twice, with no experimental manipulation. The first sample completed the second IAT after a week, showing a test-retest correlation of $r = 0.58$. The second sample completed the second IAT after a month, yielding a correlation of $r = 0.62$. The third sample was assessed after a year of completing the first IAT, yielding a test-retest correlation of $r = 0.47$.

In the particular case of implicit self-esteem measured with the IAT, a study showed that between two administrations of the SE-IAT, with a mean length of elapsed time between measurement points of 31.23 days, correlation was $r = 0.69$, with no experimental manipulation between administrations (Bosson et al., 2000). However, another study using a SE-IAT, administered twice in the same day found a correlation of $r = 0.52$ (Greenwald & Farnham, 2000).

These results, together with the fact that rest-retest reliabilities are consistently lower than the test's internal consistency convey that the IAT is both a measure assessing a stable trait and a measure of a temporal state, and that it is expectable to find individual variability as well as stable components of implicit attitudes using the IAT (Cunningham et al., 2001). With the introduction of the improved scoring algorithm (Greenwald et al., 2003), authors claim that method variance has been practically eliminated (Schmukle & Egloff, 2004).

Studies using experimental manipulations, believed to have an effect in the implicit attitude being measured, show that implicit attitudes are malleable. As we saw above, presenting pictures of admired people who belong to a racially discriminated group before the administration of a racial IAT, reduces implicit prejudice in the expected direction (Dasgupta & Greenwald, 2001; Dasgupta, McGhee, Greenwald, & Banaji, 2000). However, manipulation can be subtler. A sample of American college

studies completed a Young/Elderly-IAT, where it is expected that most people will have an implicit preference from young over old people, presented as photographs. Indeed the baseline IAT showed this widespread tendency. But between administrations, participants were made to believe that at the end of the experimental session, they will have to complete a memory recall task, and were asked to memorise pairs of words. Half of the sample were made to memorise pairs made of the word "young" and a positive word, and pairs comprising the word "old" and a negative word (e.g. youth-happy, youth-love, elderly-filth, elderly-murder, etc.). Those participants who had to memorise elderly/positive pairs, significantly changed their time-reaction response in the second IAT, showing a preference for old people. Participants allocated to the young/positive group did not change their scores. Measures of explicit attitudes towards young and old people did not change as a result of the experimental manipulation (Karpinski & Hilton, 2001).

Regarding implicit self-esteem, as we have reviewed in the previous chapter, it has been proven that implicit self-esteem reacts to experimentally delivered and day-to-day threats to the ego in a compensatory process in healthy people (DeHart & Pelham, 2007; Jones et al., 2002; Rudman, Dohn, & Fairchild, 2007).

More akin to the context of this chapter, experimental manipulations have also shown effects on implicit self-esteem in depressed patients. Gemar, Segal, Sagrati, and Kennedy (2001) examined implicit self-esteem of participants who were either currently depressed, formerly depressed and healthy controls. A SE-IAT was administered at baseline and after a depressive mood induction, which included a 10-minute exposure to a quite dark piece of music, widely used for this effect[9], while instructing participants to remember situations in their lives where they felt sad. After the manipulation the control group showed no significant effect of the induction on their implicit self-esteem scores, while the formerly depressed participants showed a reduction in their explicit and implicit self-esteem. The reduction in implicit self-esteem was so that it matched self-esteem scores of currently depressed patients at baseline. Depressed patients were not subject to experimental mood induction. De Raedt and colleagues (De Raedt et al., 2006) re-inspected the data obtained by Gemar's study, and observed that implicit self-esteem of both currently and formerly depressed participants was higher than in controls, and that the significant drop in implicit self-esteem of the formerly depressed

---

[9] Alexander Nevski's "Russia Under the Mongolian Yoke", reproduced at half-speed.

patients after mood induction was due to the higher level of implicit self-esteem of formerly depressed participants at baseline. Mood induction reduced this level drastically, which was reported matched that of currently depressed participants. But it also matched the level presented by the control group, whose implicit self-esteem levels remained stable after mood induction. This same team of researchers replicated this result (Franck et al., 2008). They found a similar implicit self-esteem level between currently depressed participants and controls, and a significantly higher levels in formerly depressed patients (who were not currently depressed because of past intervention). As in Gemar's study, after the same negative mood manipulation process, formerly depressed patients showed a reduction in implicit self-esteem that matched the level of never depressed controls. At baseline, currently and formerly depressed individuals showed a specific discrepancy with higher implicit than explicit scores. At baseline, formerly depressed participants showed similar levels of explicit self-esteem than controls, but much higher implicit self-esteem. Currently depressed patients showed a similar level of implicit self-esteem than never depressed controls, but a significantly reduced explicit self-esteem level. This group of experiments not only show that implicit self-esteem is reactive in the context of depression, but it also makes licit to surmise that interventions for depression must have increased both implicit and explicit self-esteem, when considering the baseline self-esteem status of currently and formerly depressed participants.

However, a study using a Depression IAT (self-associations with words related to the poles happy/sad), found that after the same mood induction of the other two studies, only healthy control showed a decrease in the associations between self and happy, while there was no change in formerly depressed participants (Meites, Deveney, Steele, Holmes, & Pizzagalli, 2008).

Regrettably, these few studies about change in implicit self-esteem as result of intervention did not include intervention or manipulation on currently depressed patients, whose baseline scores were used only as a benchmark. From the studies reviewed in the previous chapter we can conjecture than in these studies, currently depressed patients presented a damaged self-esteem, with implicit self-esteem higher than explicit. Formerly depressed patients had the same discrepancy, but their higher levels of explicit self-esteem were comparable to people who have never been depressed. However, in people with remitted depression, the risks associated to damaged self-esteem remain, particularly for internalising disorders (Creemers, 2014; Creemers et al., 2012; Creemers et al., 2013; Kesting et al., 2011; Pavlickova et al.,

2014), which makes sense in the light of depression being a highly recurrent disorder (American Psychiatric Association [APA], 2013; Lewinsohn, Hoberman, & Rosenbaum, 1988). But at the same time, in participants with remitted depression, higher implicit self-esteem might be playing a protective role: as we have seen, people with this kind of self-esteem discrepancy are more prone to interpret neutral information positively and seem to pay more attention to positive than negative feedback (Schröder- Abé et al., 2007b). In the experimental manipulation studies cited above, participants were explicitly asked to feel sad. But maybe, in the absence of an explicit instruction, formerly depressed patients might have perceived the mood induction paradigm as more positive than controls or currently depressed participants. Also we must consider that people with low explicit self-esteem tend to react to threats to the ego with a decrease in implicit self-esteem (while people with high explicit self-esteem react with an increase) (Jones et al., 2002). Therefore we can also surmise that formerly depressed patients, regardless of having achieved a high level of explicit self-esteem, are still inclined to react to threats in the same way than depressed (low explicit self-esteem) people, i.e. with a decrease in implicit self-esteem, compromising their strategies for stress regulation.

All these studies show that implicit self-esteem is malleable, but that at the same time represents a stable personal trait, and that variations in time must be found within subjects. These studies have utilised both experimental manipulations or taken advantage of day-to-day events to assess the responsiveness of this trait. Regrettably there is a disconcerting paucity of studies addressing changes in implicit self-esteem as a result of psychotherapy. And there are no studies, to my knowledge, addressing change in self-esteem discrepancy as result of psychotherapy.

It is disconcerting, considering that one of the first studies using the IAT was in the context of psychotherapeutic intervention. Thirty-one participants with spider phobia were offered a three-session group therapy using gradual in-vivo exposure (Teachman & Woody, 2003). An implicit associations test measured the implicit relationship between the pair "snake/spider" and the pair "afraid/unafraid", and another IAT measured the association between the pair "snake/spider" with the pair "disgusting/appealing" before and after treatment. Explicit measures of fear and avoidance showed the effectiveness of the treatment. The IATs showed significant differences between phobic patients and healthy controls at baseline, and a significant decrease in negative automatic associations regarding spiders at the end of treatment. However, a study with arachnophobic patients found no differences in a threat and

disgust IAT after 2.5 hours of exposure therapy, while explicit measures of aversion and phobia did decrease (Huijding & de Jong, 2007).

A small study (N = 30) with patients presenting a specific phobia to blood-injections-injury measured the efficacy of a session of 30 minutes of exposure therapy, explicit tests showed a significant decrease in phobic symptomatology at post-treatment which was further reduced at 1-week follow-up. However, an IAT measuring associations between images of mutilations or neutral images to the value pair "appealing/disgusting" did not change significantly at post-treatment or follow-up (Covington, 2014).

A nonrandomised controlled study of a cognitive-behavioural intervention for social phobia, of a mean length of 10 hours, showed significant changes in implicit anxiety measured by an Anxiety-IAT for socially anxious people who underwent treatment. Besides this within group change, post-treatment implicit anxiety was not different than scores obtained by the healthy control group. At baseline, the socially anxious group presented increased implicit anxiety. The healthy group's D scores remained stable (Gamer et al., 2008).

Given that in depression, as in other mental disorders, negative thoughts and feelings are experienced as mental events, rather than as the self, some authors have conjectured that it is the lack of awareness about one's own automatic reactions which is a marker of risk for depression (Teasdale et al., 2002). The success of a psychotherapy in preventing relapse will then necessitate a therapy that can increase awareness of such automatic reactions. Therapies based on awareness, like mindfulness- and mentalization-based treatments should then be able to enhance the recognition of automatic reactions and modify them, reducing risk for relapse (Phillips et al., 2010). However, the empirical literature for these therapies does not hitherto include measures of implicit attitudes. The study is therefore the first one assessing implicit attitude change as result of a psychotherapy including an awareness component, in this case Dynamic Interpersonal Therapy, which is based on the theories of attachment and mentalizing.

Hypothesis

Given this brief revision of the literature and considering the results obtained in the previous studies in this thesis, it is expected that implicit self-esteem scores will not

predict levels of depression, and therefore will not be affected by treatment group. Implicit self-esteem should not be directly related to symptomatic outcomes.

It is also expectable that implicit self-esteem scores will show a stable level along time, presenting good test-retest indices. It is then logically expected that baseline implicit self-esteem scores will not predict outcomes at 16 weeks. On the same line, changes in implicit self-esteem scores from baseline will not be related to symptomatic outcomes at 16 weeks, nor to the change in symptomatic scores from baseline to 16 weeks.

Self-esteem discrepancy, operationalised as the difference between explicit and implicit self-esteem scores, will predict measures of depression and suicide. It will also predict other symptomatic outcomes, particularly internalising disorders.

Method

Design

This is a randomised controlled trial design, phase II. Participants in two clinical sites were randomly allocated to an intervention or to a control group (minimal treatment provision on an enhanced waiting list condition) and measured at 3 time points: baseline, 8 weeks follow-up (mid-treatment) and 16 weeks follow-up (end of treatment). The present study is an analysis of the Self-Esteem Implicit Association Test at these three time points and its relationships with time, experimental group allocation and symptomatic measures.

Procedure

Patients were identified by the usual triage units of the two clinical sites. If fulfilling criteria for major depressive disorder and considered suitable for high intensity treatment were invited to participate of the study. Those who consented were randomly allocated to either weekly Dynamic Interpersonal Therapy (DIT) or an 16-week enhanced waiting list. Both groups were assessed at baseline, at mid-treatment (8 weeks) and at the end of treatment (16 weeks).

Participants

The original aim of the REDIT trial was to randomise 100 consecutive patients with a diagnosis of major depressive disorder to either condition. Inclusion criteria were

18 years of age or older, diagnosis of major depressive disorder with or without dysthymic disorder and a HDRS score greater than 14. Exclusion criteria include current psychotic symptoms or bipolar disorder, current use of antipsychotic medication, severe personality disorder, historic or current self-injury/parasuicide, historical or current primary eating disorder, current excessive use of drugs and/or alcohol, participation in another depression clinical trial within the last year where the participant has received cognitive-behavioural therapy (CBT), previous unsuccessful CBT treatment, clinical contraindications to short-term psychotherapy, evidence of pervasive use of help or highly unstable/insecure life arrangements.

Measures

Participants were assessed at baseline, 8 weeks and 16 weeks with the following assessment battery, listed in order of administration.

Hamilton Depression Rating Scale (HDRS-17)

This is the primary outcome measure of the REDIT study. Described in Chapter 4, this is a semi-structured interview assessing the presence and severity of depression. It is a widely used measure in research and clinical settings, and his psychometric properties range from adequate to excellent. (Bagby et al., 2014; Hamilton, 1960). This test yields a total score, of which larger values imply increased severity of depression.

Self-Esteem Implicit Association Test (SE-IAT)

The same iPad version used throughout this thesis was also administered to both patients and controls at baseline, mid-treatment and end of treatment. For a detailed description of this measure, see Chapter 3.

Beck Depression Inventory, Second Edition (BDI-II)I

Also described with more detail in Chapter 3, this is a 21-item self-report assessing depressive symptomatology. Participants respond these items in a 0 to 3 scale, where increased scores reflect more severity of symptoms. It yields a total score ranging 0-63 (Beck et al., 1996; Dozois et al., 1998).

EuroQoL (EQ-5D)

Developed by the EuroQol Foundation[10], this is a widely used survey instrument for measuring economic preferences for health states. It includes five items, scored 1-3, in the following health dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. These 5 items are transformed into a health index score ranging from 0 to 1, where 0 represents full health and 1 represent death or even a state regarded as "worse than death" (Hurst, Kind, Ruta, Hunter, & Stubbings, 1997, p. 71). The measure also has a 20 cm Visual Analogue Scale (VAS), where participants are required to quantify the quality of their health status, on the day of completing the test, ranging from 0 ("Your worst imaginable health state) to 100 ("Your best imaginable health state). Intraclass correlation coefficients range between 0.41 and 0.89 in various studies (Hunger, Sabariego, Stollenwerk, Cieza, & Leidl, 2012; Sonntag et al., 2013; Stark, Reitmeir, Leidl, & König, 2010) with different populations.

The Social Adjustment Scale – Self Report (SAS)

This is a 45-item self-report questionnaire measuring adjustment to daily social interactions. It assesses instrumental and expressive role performance over the past two weeks in seven different social domains: work as a worker, household tasks (for individuals whose main work is household maintenance), social and leisure activities, problems with extended family, marital problems, problems with parental role, and problems with nuclear family. Each one of these dimensions is evaluated in four major categories: the participant's performance in expected tasks, the amount of friction with others, finer aspects of interpersonal relationships, and inner feelings and satisfactions (Weissman, Prusoff, Thompson, Harding, & Myers, 1978). Each items is rated on a 5-point Likert type scale, with larger scores representing more impairment. It yields one average score for each domain and an overall social impairment score. Averages of the internal consistency of the various subscales have been reported to range from 0.61 to 0.73 (Suzuki et al., 2003; Zweig & Turkel, 2007), and test-retest reliability has been reported to be 0.80 (McDowell, 2006).

The Brief Symptom Inventory (BSI)

This 53-item self-report questionnaire has been introduced in the previous chapter. It yields nine subscores corresponding to various primary psychopathological

---

[10] http://www.euroqol.org/

dimensions and three global indices of general psychopathology and distress (Derogatis, 1993; Derogatis & Melisaratos, 1983). The total score is the General Severity Index (BSI-GSI).

Experiences in Close Relationships Questionnaire – Revised (ECR-R)

This is a 36-item self-report questionnaire measuring levels of anxiety and avoidance in romantic attachment relationships (Fraley, Waller, & Brennan, 2000). This revised version applies item-response theory to the original ECR (Brennan, Clark, & Shaver, 1998), and its internal consistency has been typically found to be > 0.90. It yields two scores, one for avoidance (discomfort with intimacy) and another for attachment anxiety (fear of rejection and abandonment). This measure was completed by participants at baseline and 16 week follow-up only.

Data Analysis

Data was retrieved from POD in Microsoft Excel format, where few missing values were imputed using expectation-maximisation method. *D* scores were calculated as in previous chapters (Appendix I), where higher scores represent higher implicit self-esteem. Data was analysed using both SPSS 21 and STATA 14 for Macintosh.

Group differences will be assessed by t-tests or Mann-Whitney U tests according to variable distribution. If *D* scores appear related to demographic characteristics of the sample, these demographics will be included in subsequent analyses.

Given that participants completed measures at 3 different time points, all regression analyses on the full sample will allow for random intercepts by participant, especially considering that we should expect individual variability between administrations (Schmukle & Egloff, 2004). When one of these regression models does not gain in predictive power by allowing for a random intercept, this will be reported. An unstructured covariance matrix was used, being it the most conservative. These mixed-effects linear and ordered logistic regression models will be utilised to ascertain the incremental predictive power of *D* scores to the predictive power of time on outcome measures, both ordinal and continuous. Similar regression models, but without defining random covariance, will be used to predict symptomatic measures at 16 weeks by baseline *D* scores, and by the amount of change of *D*, calculated as the difference between *D* scores at baseline and end of the treatment. The same process will attempt at predicting change in symptomatic measures, calculated as the difference between outcome scores at baseline and at the end of treatment. In a further step to each of these

regressions, group will be entered as covariate, and if there is an effect of that last variable, it will be reported. This whole process will be also carried out using self-esteem discrepancy values as predictors, as in the previous chapters. Discrepancy in self-esteem was calculated as the difference between the mean of the reversed and standardised BDI self-esteem items minus the standardised D score. These discrepancy scores will not be contrasted with BDI-II scores to avoid collinearity issues. Implicit self-esteem scores and the explicit self-esteem composite will also be entered as factorial predictors in these regression models, to assess the interaction of explicit and implicit self-esteem regardless of their discrepancy.

Results

The composition of the final sample at the three time points is shown in Table 27. The mean of age at baseline was 38.46 years for the control group and 37.06 for the DIT group.

Spearman-Brown coefficient for split-half reliability on difference scores of odd trials against even trials for this SE-IAT was 0.888 for baseline, 0.820 for at 8 weeks, and 0.911 at 16 weeks.

At baseline the experimental and control groups were matched according to age ($t = 0.912$, $p = 0.364$), gender ($\chi^2 = 0.927$, $p = 0.336$), baseline depression levels ($t = 0.256$, $p = 0.798$ for the BDI-II), and D score ($t = -0.882$, $p = 0.381$). There were no differences in D score between genders at each time point ($U = 861.00$, 343.50, 87.00, all $p > 0.05$, respectively for each time point). D scores were not related to age at any time point. Also, D scores were similar between intervention groups at all time points ($t_{76} = -0.882$, $p = 0.38$ for baseline; standardised Mann Whitney U statistic for 8 weeks: $z = -0.433$, $p = 0.66$; $z = 0.761$, $p = 0.44$ for 16 weeks). D scores did not change as a factor of time, as shown by a non-significant mixed-effects linear regression with random intercepts by participant ($\chi^2 = 0.49$, $p = 0.485$).

As a measure of test-retest reliability, nonparametric correlations for D scores at different time points were calculated. Eight-week correlations were $rho = 0.779$, $p < 0.001$ for baseline and 8 weeks, and $rho = 0.838$, $p < 0.001$ for 8 to 16 weeks.

Table 27: Descriptive statistics for this sample according to time point and group

		Baseline (N = 107)	**8 weeks (N = 83)**	**16 weeks (N = 79)**
Gender (M/F)	Control	20/34	16/25	14/24
	DIT	15/38	12/30	12/29
HDRS (Mean/SD)	Control	19.44/4.19	15.00/6.90	14.32/7.13
	DIT	18.85/4.27	15.20/6.06	10.80/7.01
BDI-II (Mean/SD)	Control	33.39/8.97	26.54/12.93	25.60/12.45
	DIT	33.49/7.84	26.00/11.68	18.17/13.15
BSI-GSI (Mean/SD)	Control	1.66/0.56	1.34/0.79	1.21/0.71
	DIT	1.67/0.66	1.36/0.72	0.98/0.77
EQ-5D (Mean/SD)	Control	55.38/18.66	62.10/18.58	59.24/19.48
	DIT	59.16/16.11	66.36/13.94	71.25/14.17
SAS (Mean/SD)	Control	2.73/0.37	2.51/0.50	2.46/0.45
	DIT	2.73/0.40	2.52/0.43	2.20/0.50
IAT *D* (Mean/SD)	Control	-0.256/0.72	-0.070/0.74	0.040/0.66
	DIT	0.116/0.70	-0.001/0.71	-0.123/0.72

D Scores and their Relationship with Outcome Measures

Likelihood-ratio tests on hierarchical mixed-effects linear regressions with random intercept by participant showed that *D* was not a predictor, above time, of BSI-GSI ($\chi^2 = 2.47$, $p = 0.116$), EQ-5-D index score ($\chi^2 = 2.60$, $p = 0.272$), EQ-5-D continuous score ($\chi^2 = 0.83$, $p = 0.661$) and SAS overall score ($\chi^2 = 0.61$, $p = 0.738$).

Regarding the relationship of implicit self-esteem and depression mixed effects linear regressions, with random intercepts by participant showed that *D* scores do not add to the predictive power of time on total scores of either BDI-II or HDRS for any of the groups (intervention or control), as shown by log-likelihood ratio tests comparing a model with time as only predictor and a second model with *D* entered as covariate ($\chi^2 = 2.80$, $p = 0.246$ for the BDI-II; $\chi^2 = 3.80$, $p = 0.150$ for the HDRS). The addition of group allocation as a predictor did not improve the predictive model from one with time as unique predictor ($\chi^2 = 3.82$, $p = 0.282$ for the BDI-II; $\chi^2 = 4.01$, $p = 0.260$ for the HDRS).

For the prediction of levels of severity in the full sample, sequential mixed-effects ordinal logistic models, with random intercepts by participant showed that D scores are significant to predict the odds of being diagnosed as severely depressed. For severity of depression assessed by the BDI-II, the increase in one unit of D was associated with an increase in the odds of being diagnosed as severely depressed versus less severe categories of depression, with an odds ratio of 104.278 (95% CI, 6.611 to 1644.799) at 8 weeks, and OR = 37.717 (95% CI = 2.131 to 667.477) at 16 weeks, when explicit self-esteem and intervention group are kept constant. This model was generally significant ($\chi^2 = 43.95$, $p < 0.001$) and it is a better fitting model than one containing time, explicit self-esteem an intervention group as predictors ($\chi^2 = 30.51$, $p < 0.001$). The same model failed to find significant results for severity measured by the HDRS.

A hierarchical linear regression was performed, to predict BDI-II total scores at 16 weeks by baseline D score. The first model included explicit self-esteem and group allocation as predictors. The addition of implicit self-esteem to the model did not achieve a significant improvement in model fit, as assessed by a likelihood-ratio test ($\chi^2 = 0.21$, $p = 0.650$). A similar model was fit to the data, but this time the second model entered change in D scores from baseline to 16 weeks. The second model was, overall, a better fit to the data ($\chi^2 = 4.12$, $p = 0.043$). However, the regression coefficient for baseline D scores was not significant ($\beta = 5.283$, $p = 0.063$). This same process was carried out to predict total HDRS scores. The addition of baseline D scores did not yield a better fit for the data than a model including explicit self-esteem and group allocation as predictors ($\chi^2 = 1.84$, $p = 0.174$). A model adding change in D scores from baseline to 16 weeks was not significant ($F_{3,24} = 1.91$, $p = 0.155$).

Change in symptomatic measures from baseline to end of treatment was the dependent variable for hierarchical linear regressions aiming at assessing the predictive power on this measure of change by D scores at baseline, by change in D scores from baseline to 16 weeks, and by their interaction. The last step included group allocation as predictor. None of these models was statistically significant in the prediction of HDRS or BDI-II change scores. Models were also non-significant in the prediction of change in BSI scores, of change in EQ-5-D scale scores and for the change in SAS scores.

With regards to the relationship between D scores and suicide ideation, a mixed-effects ordinal logistic regression with random intercepts by participant and with time and implicit self-esteem as predictors of suicidal ideation did not achieve statistical significance ($\chi^2 = 6.46$, $p = 0.091$).

None of the regression models described in this section gained in predictive power when allocation group was added as predictor.

Self-esteem discrepancy

The index for self-esteem discrepancy was calculated using four BDI-II items thought to reflect explicit self-esteem, namely feeling like a failure, self-dislike, self-criticalness and worthlessness. Therefore, discrepancy scores will not be contrasted against BDI-II scores. Explicit self-esteem scores were not related to D scores (*rho* = 0.075, $p = 0.349$). Explicit self-esteem scores did not change as function of either time, condition or their interaction. The discrepancy scores obtained were also not related to time and group allocation.

Regarding HDRS total score, sequential mixed-effect linear regressions showed that the addition of self-esteem discrepancy to a model with only time as predictor was a better fit for the data ($\chi^2 = 32.78$, $p < 0.001$). A decrease of one unit in self-esteem discrepancy (implying a relative increase of implicit self-esteem over explicit self-esteem) significantly predicts a decrease in HDRS total score at 8 weeks ($\beta = -2.557$, $p = 0.001$) and 16 weeks ($\beta = -2.093$, $p = 0.027$), as shown in Figure 16. The overall effect of the interaction between time and self-esteem discrepancy scores was significant ($\chi^2 = 11.24$, $p = 0.004$).

Figure 16: Fixed-portion of mixed-effect prediction of HDRS total scores by self-esteem discrepancy, according to time point. (ns)= non-significant.
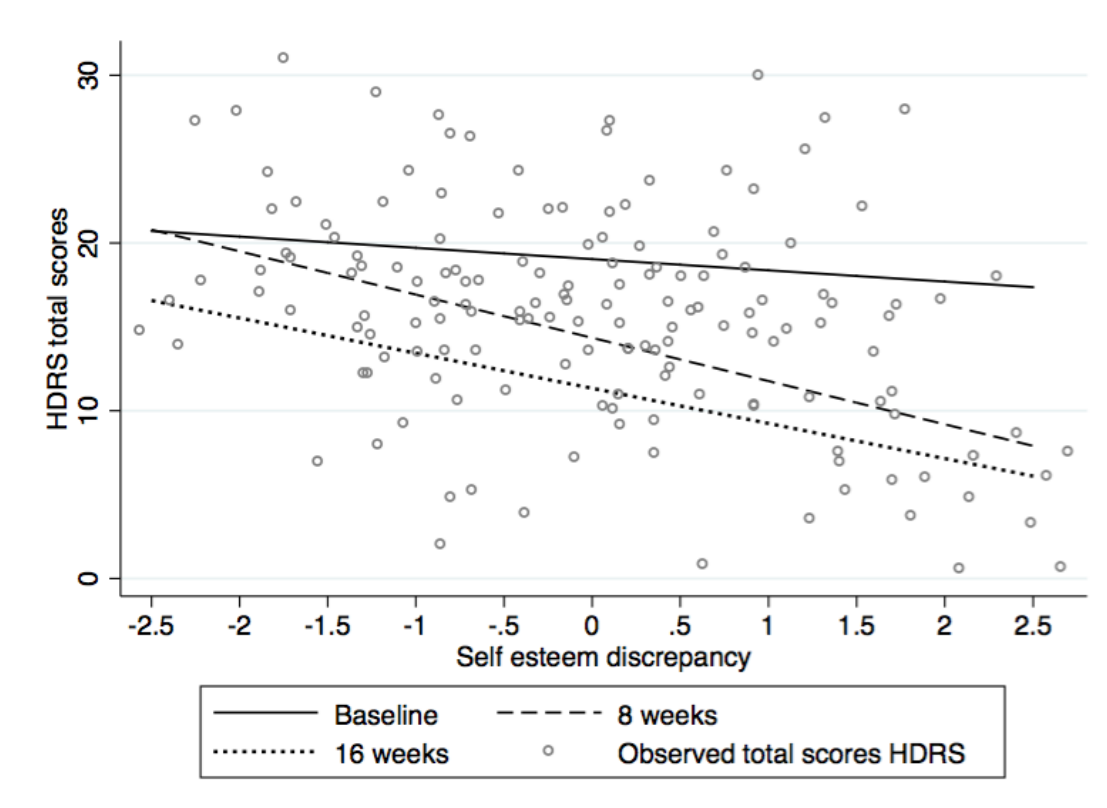
Self-esteem discrepancy scores were also significant predictors of the BSI-GSI in mixed-effects linear regressions with random intercepts by participant. Coefficients were significant for baseline scores (β = -0.146, p = 0.010), but non-significant for 8 weeks (β = -0.127, p > 0.050), and 16weeks (β = -0.110, p = 0.124). The overall interaction was not significant (χ^2 = 4.24, p < 0.120). However, when time is kept constant, self-esteem discrepancy scores have a significant main effect (β = -0.223, p < 0.001). Once again, the direction of the regression indicates that general severity of psychopathology measured by the BSI increases when self-esteem discrepancy decreases, which implies a relative increase in implicit self-esteem.

For the somatic subscale of the BSI, there was no interaction between time and self-esteem discrepancy, but a multiple regression model with random intercepts by participant and random slopes by time revealed significant main effects for time (β = -0.154, p = 0.010) and discrepancy (β = -0.189, p = 0.002). For the obsessive-compulsive subscale of the BSI, self-esteem discrepancy was a significant predictor when controlling for time (β = -0.338, p < 0.001). Furthermore, the prediction of this subscale's scores was significant for the interaction between explicit and implicit self-esteem (β = -0.342, p = 0.010), as shown in Figure 17, showing that people with damaged self-esteem are more likely to show more obsessive-compulsive problems. Self-esteem discrepancy was a significant predictor, when controlling for time, of BSI depressive subscale (β = -0.298, p < 0.001), anxiety subscale (β = -0.270, p < 0.001), psychoticism (β = -0.196, p < 0.001) and phobic anxiety (β = -0.239, p < 0.001). This last outcome was also predicted by a significant interaction between explicit and implicit self-esteem (β = -0.285, p = 0.023). Suicide, measured by the HDRS, was also predicted by self-esteem discrepancy, as assessed by a mixed effects ordinal logistic regression with random intercepts by participant. When controlling for time, an increase

Figure 17: Fixed portion of the interaction of explicit and implicit self-esteem in the linear prediction of BDI obsessive-compulsive scores
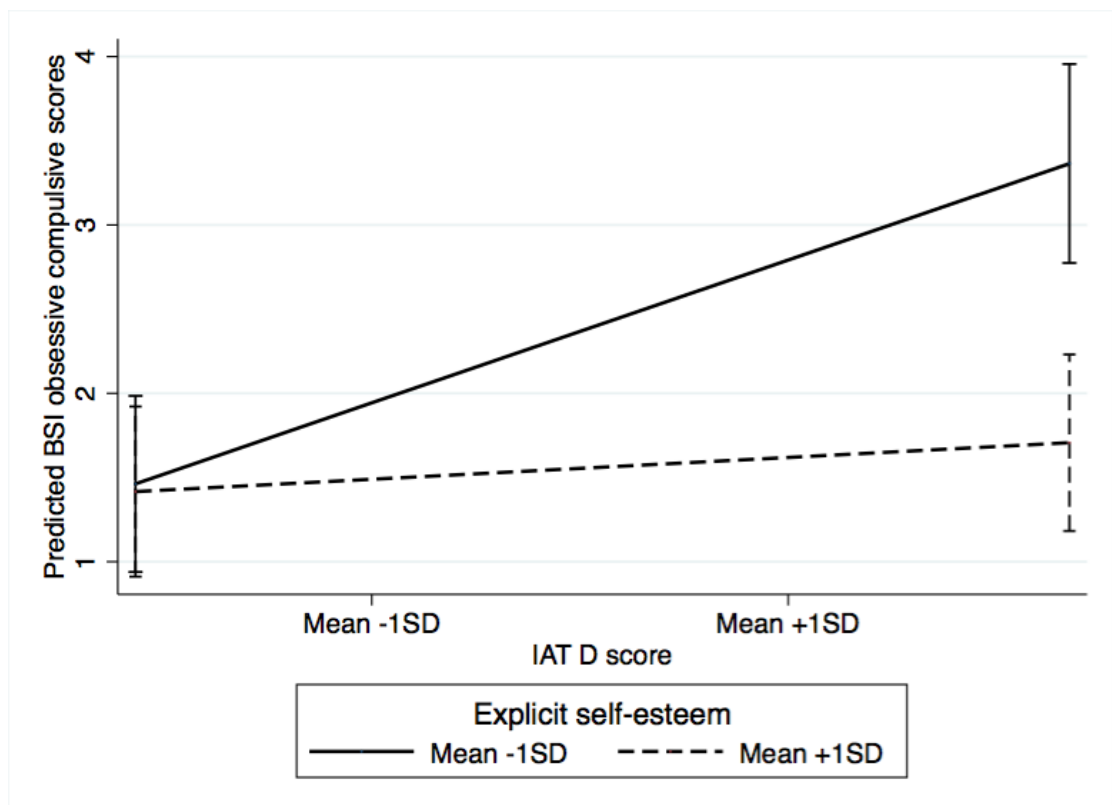
of one unit of self-esteem discrepancy the odds of presenting less severe suicidal ideation are OR = 0.50 times greater than presenting more severe suicidal ideation, keeping time constant in the model ($p = 0.005$).

The VAS general health index of the EQ-5-D, a measure of subjective health status ranging from 0 to 100, was not predicted by self-esteem discrepancy. However, a mixed-effects linear regression with random intercepts by participant showed a significant interaction between explicit and implicit self-esteem in the prediction of this general health index, when controlling for time and group allocation ($\beta = 5.080$, $p = 0.048$), as shown in Figure 18. Participants with high explicit and implicit self-esteem levels show more positive scores in subjective health status, as well as people with low explicit and implicit self-esteem. People with discrepant self-esteem, regardless of the direction of such discrepancy tend to report more negative health.

Explicit, implicit self-esteem scores and their discrepancy were not related to measures of attachment assessed by the ECR.

Figure 18: Fixed portion of the interaction between explicit and implicit self-esteem in the linear prediction of EQ-5-D health state

155
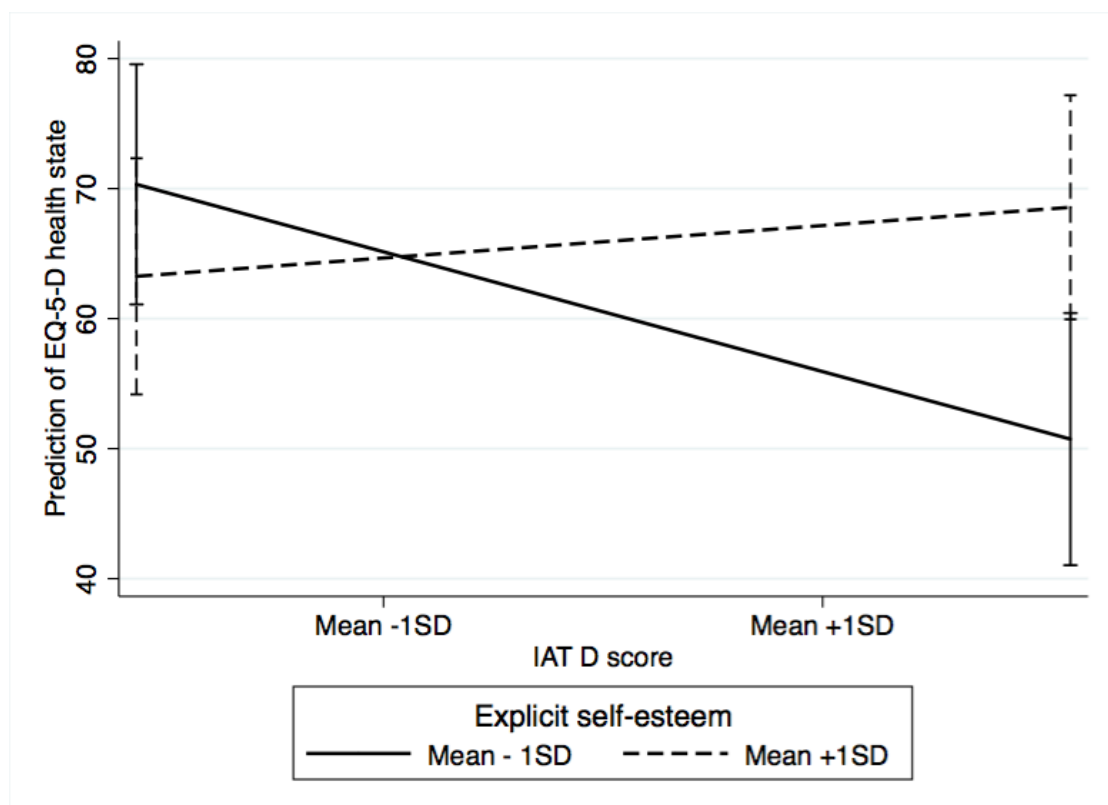
Discussion

This last empirical chapter analysed the SE-IAT's relationships with a sample of adult depressed patients randomised to a psychodynamic psychotherapy for depression or to a control group consisting on an enhanced waiting list. Both conditions had a duration of 16 weeks. In particular, the change in implicit self-esteem as a function of time and intervention was assessed, as well as the predictive power of implicit self-esteem, and the relationship between D scores and changes in other symptomatic scores. Besides implicit self-esteem D scores, an explicit self-esteem composite was calculated from the BDI-II, and the discrepancy between this score and D scores were analysed in conjunction with symptomatic measures.

In line with the hypothesis, D scores did not change after psychotherapy in comparison with the control group. Furthermore, none of the groups showed significant changes in D scores as function of time. In fact, correlations between D scores obtained at each time point, taken as test-retest reliability for this version of the IAT, ranged between acceptable and good. Previous chapter have shown that D scores are unable to discriminate between depressed and non-depressed individuals, so it was expected that psychotherapeutic intervention would not cause changes in implicit self-esteem scores, regardless of depression remission. If we compare the 8-week test-retest reliability for this IAT (mean rho = 0.809) with Nosek's test-retest reliability formula mentioned in the introduction to this chapter (Nosek et al., 2007a), the present SE-IAT was much more robust to the passage of time and to the influence of intervention.

Furthermore, analyses of the relationship of D scores in the full sample were showed that group allocation was not a concurrent predictor, with D and time, of symptomatic scores. Implicit self-esteem did not predict, above time, general symptomatic severity, social adjustment and health status. Implicit self-esteem was not related to measures of depression. It is important to note here that depression scores were not predicted by group allocation, above time.

However, D scores were related to severity of depression only in the BDI-II. However, this relationship was inverse to that observed in Chapter 4. In this sample, an increase in implicit self-esteem was associated with the increase in the odds of presenting more severe levels of depression. Differently from Chapter 4, implicit self-esteem was not related to suicidality levels.

Baseline implicit self-esteem did not predict symptomatic scores at the end of treatment, and changes in implicit self-esteem from baseline to end of treatment also failed to predict symptomatic scores in all measures at 16 weeks.

Regarding the relationship between implicit self-esteem and suicide, D scores were unrelated to measures of suicidality, differently from the findings in Chapter 4.

Taken together, these results confirm that implicit self-esteem scores, measured with the IAT, are not related to measures of depression when D is considered as an independent predictor or, in this case, in relationship with time and the presence of intervention.

As in the previous chapter, interesting results emerged from the discrepancy between explicit and implicit self-esteem. It is important to notice here that explicit self-esteem was also unrelated to treatment condition and time, which precludes that the significant results found for self-esteem discrepancy correspond uniquely to changes in explicit self-esteem. However, considering the strong correlations in implicit self-esteem between time points (and therefore its stability), we can surmise that a great portion of the variability in discrepancy scores is indeed motivated by changes in explicit self-esteem while implicit self-esteem remains relatively stable.

Self-esteem discrepancy, understood as the difference between explicit and implicit self-esteem predicted a series of symptomatic outcomes, all of which pertain to the domain of internalising problems, except for scores of obsessive-compulsive symptomatology (Caspi et al., 2014). Regarding this last outcome, it might be possible that the prediction model achieved significance with regards to the internalising aspect of the obsessive-compulsive (obsession) rather than its externalising manifestations (compulsion). The internalising problems predicted by the discrepancy between explicit and implicit self-esteem were depression, general symptom severity, psychosomatic symptoms, general anxiety and phobic anxiety. The odds of presenting with more severe suicidal ideation were also predicted by a decrease in discrepancy scores, as in the previous chapter and in the literature (Creemers et al., 2012).

All these outcomes predicted by self-esteem discrepancy showed the same direction. Symptomatic level increases when discrepancy decreases, indicating a relative higher implicit self-esteem over explicit self-esteem. There is then a tendency towards a damaged self-esteem style in more severely ill patients. It is not new that people who show a low explicit-self-esteem are more likely to present mental disorders (Sowislo & Orth, 2013), but the psychopathological effects of low explicit self-esteem

are not restricted to internalising problems (Donnellan, Trzesniewski, Robins, Moffitt, & Caspi, 2005; Leary, Schreindorfer, & Haupt, 1995). However, results obtained using self-esteem discrepancy were almost exclusively internalising. What discrepancy shows is that there is that decrements in explicit self-esteem are potentially risky for general psychopathology, but it is the relative level of implicit self-esteem which might be one of the factors defining which broad kind of psychopathology is the one encountered. There are in fact three cases in which the discrepancy in self-esteem levels is a predictor for internalising problems: explicit self-esteem decreases and implicit self-esteem increases, explicit self-esteem decreases and implicit self-esteem remains stable or implicit self-esteem increases while explicit self-esteem remains stable. The discrepancy obtained by any of these three cases indicates risk for internalising problems. This is then a conceptual replication of the findings by Creemers et al. (Creemers, 2014; 2013), who found this same tendency in young healthy women, using both the SE-IAT and the Name Letter Task. This is a very interesting result when thinking of the usefulness of this IAT for the study of psychopathology. The predictive value of implicit self-esteem seems to depend on its conjunction with explicit self-esteem.

It is a recent scientific certainty that risk for all kinds of psychopathology can be subsumed under a unique third-level factor, called "factor p". However, at a second level, psychopathological problems are grouped in externalising, internalising and thought disorders (Caspi et al., 2014). It is at that level that the usefulness of the SE-IAT in depressed patients seems to lie. It contributes to differential diagnosis and helps explicate mechanisms of the "choice of neurosis", a topic on clinical psychology as ancient as clinical psychology itself (Freud, 1896). Theoretical explanations for this phenomenon were revised in the previous chapter.

Regarding the goals of this chapter, and the overall thesis, it has been found that the SE-IAT, when related to explicit self-esteem scores, gives more specificity to the diagnosis of depression, predicting that the tendency towards damaged self-esteem predicts internalising problems.

Limitations

One of the most important limitations of this report is the fact that it was elaborated previous to ascertaining the effectiveness of the treatment. The absence of change in explicit self-esteem seems to point to a treatment of limited effectiveness.

That is why most implicit self-esteem effects have been analysed in the full sample. The results on the effectiveness of DIT in comparison with the control will be reported elsewhere (Fonagy et al., in preparation).

It is also a limitation to this study, as well as the previous ones, not to have counted with a validated measure of explicit self-esteem. The self-esteem composite created for this study is a reliable one, but it is probably less nuanced to the different aspects of explicit self-esteem.

The objective of this thesis was to assess the pertinence and usefulness of the Implicit Association Test (IAT) in different interpersonal domains, particularly in the treatment of depression. Five empirical studies on different samples are reported. For four of these five empirical studies, an adaptation of the Self-Esteem Implicit Association Test was developed for administering via portable electronic tablets. Results were mixed, as is observable in the extant literature for this measure, particularly when it is applied to clinical investigations.

Introduction: A Scientific Field in Need for Replication

This history of inconsistency of results was what provided the motivation for this thesis in the first place. The rationale behind the IAT is convincing and theoretically sound, and consistent and useful results have been found when it has been applied in the field of social psychology and market research[11]. For example when studying consumer behaviour, IATs comparing implicit preference for one brand over another are strongly correlated with explicit attitudes towards the target brands. Such discovery would not have any usefulness, if we consider the explicit attitudes alone. However, implicit attitudes measured by the IAT are capable of predicting consumer behaviour and product usage above and beyond what can be predicted from explicit attitudes (Brunel, Tietje, & Greenwald, 2004; Maison et al., 2001; Maison, Greenwald, & Bruin, 2004).

The IAT also shows robust results in social psychology: a universal in-group preference when it comes to, for example, race (Baron & Banaji, 2006; Cunningham et al., 2004), gender stereotypes (Greenwald & Nosek, 2001; Rudman & Goodwin, 2004), nationality, religion and attitudes towards old age has been found (Rudman, Greenwald, Mellott, & Schwartz, 1999). In general it is possible to see that the implicit association test is capable of dodging self-presentational biases and social desirability in topics that are socially controversial. Studies in this field have also revealed the variability of certain IATs: white participants' implicit preference for white over black people decreases if before the IAT participants are presented with photos of admired black people (Dasgupta & Greenwald, 2001). And in the case of findings from market

---

[11] For a full list of social psychology and market research topics where the IAT has been used, see (Nosek & Smyth, 2007, pp. 16-17)

research, implicit racial attitudes measured by the IAT have also been shown to predict behaviour (McConnell & Leibold, 2001). However, unlike in the findings from market research in this area, in social psychology studies the relationship between implicit and explicit attitudes is rarely correlated due to the effect of the desire to mask shared cultural stereotypes that are unacceptable (Arkes & Tetlock, 2004; Kiefer & Sekaquaptewa, 2007; Yamaguchi et al., 2007).

In sum, it is possible to find different relationships between explicit and implicit measures depending on the motivation for positive self-presentation. This is further confirmed by studies that show that participants completing self-report measures while subject to distractions or other types of cognitive load will present an increased correspondence between implicit and explicit measures (Hofmann et al., 2005a). The same can be said in relation to the capacity of explicit measures to predict behaviour. More spontaneous behaviour will be predicted better by implicit than explicit measures (Friese, Hofmann, & Schmitt, 2009).

In order to contextualise the findings of this thesis, it is necessary to briefly revise the current state of IAT measurement in the context of psychopathology. Results with the IAT are not as consistent or straightforward when it comes to measuring personality traits or psychopathology as it is in social psychology and market research. In fact, when measuring personality or psychopathology, the motivation for a socially desirable self-presentation is not warranted in the same way that it is with racism or discrimination. Although there might be some self-presentation biases at play, the IAT's goal when measuring a personality trait or a psychopathological disposition (e.g. self-esteem or fear of spiders) is to obtain a glance into automatic reactions that are, in many cases, unavailable to introspection, regardless of the moral quality of the implicit construct of interest, or its conflictive status with regards to consciously upheld beliefs (Lane et al., 2007; Nosek et al., 2007a). Therefore in some cases the implicit attitude measured is positively related to the conscious attitude, like in fear of spiders, for example(Teachman & Woody, 2003), while in others (as in the case of general anxiety) implicit and explicit measures are not correlated with each other and social desirability does not moderate this correlation (Egloff & Schmukle, 2002; Egloff & Schmukle, 2003).

In this case, implicit attitudes against spiders have an incremental predictive power to self-report via avoidance behaviour and can distinguish phobic individual from healthy controls or from individuals with a remitted phobia (Teachman et al., 2001), and

those groups from spider enthusiasts (Ellwart et al., 2006). Furthermore, brief exposure treatment had effects on both implicit and explicit arachnophobic associations (Teachman & Woody, 2003). However, as is common with psychopathological research with the IAT, these results were not replicated when controlling for general emotional response (Huijding & de Jong, 2007). There were several methodological differences between these studies that might partly explain this failure of replication (among them the duration of therapy, which is relevant given the overlearned nature of automatic associations), which illustrate that in psychopathology research the IAT is much more sensitive to other methodological factors than in social psychology research and therefore results cannot be taken at face value.

In the case of implicit anxiety, as mentioned above, IATs have shown not to be related to self-report measures of anxiety. However, implicit anxiety was related to the worsening of performance on tasks (either a cognitive task or giving a speech in public) after being subject to a stress induction, and to observer-rated anxiety. The latter index was related to self-reported anxiety, leaving implicit anxiety as only predictor of behaviour (reduction in performance during tasks) (Egloff & Schmukle, 2002). However, the same group of researchers later found that inducing stress in participants did not change implicit self-esteem scores as measured with the IAT, but was only related to self-reported anxiety. These researchers conclude that in anxiety, the IAT is a measure assessing a trait, and not a state of anxiety. They further generalised this finding by concluding that the lack of change in the IAT after manipulation has to do with the nature of the attitudinal object (Schmukle & Egloff, 2004). When the IAT measures self-associations it seems to be much more stable than when measuring implicit associations towards an entity that is not the self, such as socially discriminated groups. This was further confirmed by the temporal stability showed by the anxiety IAT in periods as long as one year (Egloff et al., 2005). This explains the change in IATs after psychotherapeutic interventions in spider phobia, where the IAT did not measure self-associations but the association of spiders to fear and disgust. But a replication of this study showed changes in implicit anxiety after group therapy for social anxiety (Gamer et al., 2008). As happens with much of the contradictory research on the IAT in psychopathology, discussion of the results of this last replication did not mention the original studies it failed to replicate. Again we see that in psychopathology, the IAT presents important variability. In fact, an IAT measuring social anxiety (instead of general anxiety) showed significant decreases after a behavioural intervention, which

were reflected in behaviour but not in self-reports of social anxiety (Clerkin & Teachman, 2010).

Regarding self-esteem, it has been repeatedly shown that people who are socially anxious show positive levels of implicit self-esteem, but these are significantly lower than the implicit self-esteem levels of non-anxious individuals (de Jong, 2002; Tanner et al., 2006). A similar study in an adolescent sample yielded the same results, but only for girls who had low explicit self-esteem. In this sample, low explicit self-esteem was common to both social anxiety and depression, but implicit self-esteem was not related to depression (De Jong et al., 2012). Another study assessed implicit self-esteem in socially anxious participants before and after inducing anxiety. Socially anxious participants showed high implicit self-esteem comparable to the control group. But after manipulation, they only showed a decrease in explicit self-esteem (Schreiber et al., 2012). The decrease in explicit self-esteem created a discrepancy between implicit and explicit self-esteem that was significantly different from the discrepancy of controls. Another study found significantly lower implicit self-esteem in socially anxious individuals than in controls, and that lower implicit self-esteem was only associated with symptom severity in men (Glashouwer et al., 2013). Another study similarly found reduced implicit self-esteem in socially anxious individuals after stress induction, but did not measure implicit self-esteem at baseline. Interestingly, people with social anxiety and comorbid depression showed lower explicit self-esteem than socially anxious patients without depression, but the difference was not reflected in implicit self-esteem (Ritter et al., 2013). Finally, a very recent study in adolescents found no relationship between implicit self-esteem and symptoms of social anxiety or symptoms of depression, and no interaction between implicit and explicit self-esteem in the prediction of symptomatic outcomes (van Tuijl et al., 2014).

In relation to the main subject of this thesis, namely the relationship between implicit self-esteem and depression, the state of the literature is similarly inconsistent. Various research designs using the SE-IAT have reached different results. A study utilising the SE-IAT, the Name Letter Preference Task and the Extrinsic Affective Simon Task, obtained two *D* scores, one subtracting negative trials related to "self" from positive trials related to self, and the other similarly subtracting positive to negative trials related to "other". Depressed patients showed more positive *D* scores relating to self than controls, warranting the conclusion that depressed patients have higher implicit self-esteem than individuals who are not depressed (De Raedt et al., 2006). Such a result appears counterintuitive, as it runs against most cognitive theories

of depression, which posit low self-esteem as one of the markers of this disorder (Beck & Alford, 1967).

However, the opposite result has also been found. A SE-IAT found that depressed people have both low explicit and implicit self-esteem, and that risk of depression was increased when people with low implicit self-esteem face adverse events in life (Steinberg, 2006; Steinberg et al., 2007). But in another study, the interaction between adverse life events and low implicit self-esteem disappeared when explicit self-esteem was added to the model as predictor (Haeffel et al., 2007). This study also applied a laboratory-based negative stressor on another sample and found that people with lower implicit self-esteem are more prone to react with depressive symptoms. Other studies have found no relationship between implicit self-esteem and depression (De Jong et al., 2012; De Raedt, Franck, Fannes, & Verstraeten, 2008; Franck et al., 2008; Lemmens et al., 2014; Pavlickova et al., 2014; Valiente et al., 2011; van Tuijl et al., 2014). Franck et al's study (2008) assessed formerly depressed individuals in remission, and found higher levels of implicit self-esteem in this subgroup compared to both controls and currently depressed patients. This subgroup showed a marked decrease only in implicit self-esteem scores after a negative mood induction. In a comprehensive meta-analysis covering the relationship between implicit self-esteem and depression, Phillips (2010) showed that implicit self-esteem is not a consistent predictor of depression when measured with the IAT. Of ten studies using this measure, three found higher implicit self-esteem in depressed individuals, and the rest found the opposite result or no relationship at all. Her own study (Phillips et al., 2012) showed significantly lower implicit self-esteem in depressed participants.

This brief introduction to the state of the evidence on the use of IAT in psychopathology, particularly in relation to depression, demonstrates the extent to which this field of study is in need of further replications. This thesis constitutes an attempt to make a contribution to the incomplete body of knowledge relating to the pertinence of the IAT in psychopathology research.

Psychology, and specifically social psychology, has been subject to debate regarding the lack of replication of results and the misleading tendency of researchers and journal to publish only positive results, or results that will attract the attention of mainstream media, while neglecting to publish negative results, with negative consequences for the advancement of science. The controversy surrounding this question has become increasingly heated since the publication in 2012 of an open e-

mail[12] from Nobel Laureate psychologist Daniel Kahneman, in which he expressed concerns about the lack of robustness of priming results, making social psychology "the poster child for doubts about the integrity of psychological research" (Kahneman, 2012, September 26, p. 1). He criticised the lack of replication of many discoveries in social psychology. This was followed by the publication of nine failed replications of a finding that was thought to be a scientific fact (Shanks, Newell, Lee, Balakrishnan, & Ekelund, 2013): people will perform better at a general knowledge test if before the test they are primed with the concept "intelligent" or the word "professor". Correspondingly, they would perform worse at the same task if primed with the concepts "stupid" or "hooligan" (Dijksterhuis & Van Knippenberg, 1998).

Daniele Fanelli has further contributed the debate, with his finding that the proportion of positive results in psychology's top journals rose by more than 22% from 1990 and 2007 (Fanelli, 2011). In fact, psychiatry and psychology are the sciences with the largest proportion of positive results in their publications: they are five time more likely to report a positive result than space sciences.(Fanelli, 2010). This disquiet is not new: as early as 1959 Sterling (1959) showed that investigations which report significance tests are more likely to remain unpublished when results are non-significant. The consequence of this is that other researchers, unaware of the existence of the previous non-significant result, might try to repeat the process and so on, until one researcher does find a significant result (by chance, Type I Error), and that results gets published, stating as a fact what in reality was a chance finding after several repetitions of the same experiment by different researchers (Sterling, 1959). Thirty years later, the same author revisited the topic and found that publication biased practices had not changed at all (Sterling, Rosenbaum, & Weinkam, 1995). The consequences of the replicability crisis are potentially devastating for psychological science, which relies on meta-analyses and other amalgamations of the literature to produce allegedly reliable knowledge (Easterbrook, Gopalan, Berlin, & Matthews, 1991).

The response to Kahneman's and Fanelli's warnings from researchers include the recently published Open Science Framework[13] collaboration project (Open Science Collaboration, 2015), a massive initiative by more than 270 psychological researchers

---

[12] http://www.nature.com/polopoly_fs/7.6716.1349271308!/suppinfoFile/Kahneman%20Letter.pdf

[13] https://osf.io/ezcuj/

around the world to replicate findings from the journals *Psychological Science*, *Journal of Personality and Social Psychology* and the *Journal of Experimental Psychology: Learning, Memory and Cognition*. Mean effect sizes of replication were $M_r = 0.197$, SD = 0.257, which was half the magnitude of original research papers; and while 97% of the original studies had statistically significant results, only 36% of the replications achieved significance. Several of these were studies on priming and other indirect techniques but none on the IAT. However, the Open Science Framework keeps working and stores information about ongoing research projects, many of them about the IAT.

Another response to this critique has been the creation of a "file drawer" website containing replications that have not been accepted for publication[14]. The only study using the IAT is a replication of a study on a sample of white women who showed increased implicit racial out-group bias when ovulating if they had previously associated black males with physical formidability (McDonald, Asher, Kerr, & Navarrete, 2011). The replication was successful (Cesario, 2014a).

As we have seen in this review of the literature, and was demonstrated throughout the thesis, the use of the IAT in clinical research is in clear need for replication. Contradictory results blur the possibility of ascertaining the usefulness of the measure, but more importantly, its underlying mechanism, which is still a matter of debate (Koole et al., 2007). In the following section, the results of the empirical studies in this thesis will be summarised, and the reasons for the success or failure of each study in replicating the findings of the current literature will be discussed.

Summary of Findings

The studies in this thesis must be understood sequentially, to attempt at an integration of the findings. The thesis starts with an introduction of the Implicit Association Test and its functioning. It starts with the broad theme of assessing the usefulness of this measure in different socio-cognitive domains, addresses its psychometric properties in different samples (including a developing sample), and finally focuses on the relation between self-esteem (as a socio-cognitive construct) and psychopathology, particularly in depression.

---

[14] http://psychfiledrawer.org/

The empirical study reported in Chapter 2 attempted to measure the influence of implicit maternal attitudes towards parenting and attachment on infant attachment style. It is the first study of this kind: the few extant studies relating attachment to the Implicit Association Test measured adult attachment in the subject who responded the IAT (Dewitte et al., 2008; Zayas & Shoda, 2004; Zayas & Shoda, 2005). Results in the study conducted in this thesis indicated that the IAT measures were unrelated to most constructs of parental style, parental psychopathology and infant attachment. Mothers' *D*-scores failed at predicting infant attachment style. Sample sizes were inadequate for the analysis of variance between 2 (secure/insecure) and 3 (secure/anxious/avoidant) groups with a power = 0.8 to find large population effect sizes with a confidence of 95%. This is because in spite of an adequate total sample size, there was a larger proportion of securely attached infants in the sample, making the insecure groups too small for reliable analysis. The study was just below the sample size to capture medium effect sizes with that same power and significance level (Cohen, 1992). This chapter was included in the thesis to help the reader attain familiarity with the versatility of the test: both of the IATs used were designed ad hoc to meet the goals of the study. The few results obtained must be interpreted with care, given that this was not a clinical population, and the results are related to psychopathological measures: mothers with an automatic preference for their infants over other infants showed an associated with increased anxiety and psychoticism and to mother-reports of their babies' distress, while mothers with an implicit preference for attachment pictures over leisure pictures showed increased levels of parental stress. The major limitation of this study, besides sample size, was the absence of an attachment measure for mothers, which could have given important information about the moderation of the relationship between parental and infant attachment by implicit attitudes. Within the context of this thesis, one of the major insights obtained from this study was the relationship between implicit attitudes and psychopathological measures, which were further explored in Chapters 4, 5, and 6.

Chapter 3 is also an introductory chapter. It introduced our new version of the SE-IAT developed specifically for this thesis. It was based on the SE-IAT described by Steinberg (Karpinski & Steinberg, 2006; Steinberg, 2006; Steinberg et al., 2007). It was administered to a normative adolescent sample with the objective of calculating its psychometric properties and its relationship with demographic variables in a demographically heterogeneous sample. It yielded good to excellent internal reliability indices ranging from 0.753 to 0.901. It was also robust to gender and age, demonstrating discriminant validity to those demographics. This study was successful

in replicating these results from the literature. There are no published accounts of demographic differences in the SE-IAT, except for one article in which this difference did not affect D scores per-se, but the relationship between implicit self-esteem, explicit self-esteem, gender and social anxiety. Socially anxious women with both low explicit and implicit self-esteem tend to show more symptoms of social anxiety (De Jong et al., 2012). There are no accounts of age differences in the completion of the SE-IAT. Finding these internal consistency levels and robustness to demographic characteristics was useful for validating this version of the IAT in this sample, which was performed on a tablet computer and involves a slightly different approach to the administration of a measure that is highly sensitive to methodological nuances. Besides changing the administration device, other changes were kept to a minimum, and all tablet computers used had the same internal clock and operative system, and stored information in a centralised repository. The results yielded by this study supported the integrity of the new version of this measure, and allowed us to administrate it to a clinical sample in subsequent chapters. The results here obtained are not to be generalised lightly. They were obtained from a normative adolescent sample, which on one side weakens the external validity of the validation process, but on the other side it ascertains its methodological robustness and the possibility of using it with younger (non-adult) samples, and still show excellent psychometric properties.

Chapter 4 is the first of a series of three studies tackling the main subject of the thesis, namely the relationship between implicit self-esteem as a socio-cognitive construct and the prevalent mental health disorder of depression. A sample of depressed participants and healthy controls were administered the tablet version of the SE-IAT and other symptomatic measures with the aim of ascertaining the relationship between implicit self-esteem and depression. It was found that D scores did not differentiate between depressed and control participants. Furthermore, D scores were positive in both groups, indicating an implicit preference for the association of positive words and the self, relative to not-me, which is in line with most of the literature showing a universal positive implicit self-esteem in different samples with different levels of psychopathology in various cultures (De Jong et al., 2012; De Raedt et al., 2006; Franck et al., 2008; Steinberg, 2006; Steinberg et al., 2007; Wegener et al., 2015; Yamaguchi et al., 2007). Sample size was marginally below adequacy for finding large effect sizes at an 80% power for a 95% confidence (Cohen, 1992). However, D scores were predictive of the probability of being diagnosed with a more severe depression, with higher implicit self-esteem lowering the probability of depression severity. Most of the

correlations found were not significant enough, but they showed a negative relationship to D scores, indicating increased symptomatic severity with lower levels of implicit self-esteem. The lack of significance is due to Bonferroni adjustments of significance values, to avoid Type I errors (a chance finding due to multiplicity of statistical tests). Decreases in D scores were related to an increase in the odds of presenting more severe suicidal ideation. These results were in line with some of the literature in depression and implicit self-esteem. However, these studies used slightly different versions of the IAT, one measuring the implicit association of self with generally pleasant or unpleasant words (like sunrise vs. vomit), not necessarily related to a self-evaluation (Conner & Barrett, 2005). Other studies show a relationship of low implicit self-esteem and negative life events in the prediction of depression in people with cognitive vulnerability, understood as higher levels of perfectionism, sensitivity to social criticism and rigid ideas about the world (Steinberg, 2006; Steinberg et al., 2007). Another study showed a similar result, but this time controlled for more factors, such as baseline depression levels, and found that the predictive power of D scores was lost when controlling for these new variables (Kruijt et al., 2013). These studies, together with the marginally insufficiently powered study reported in Chapter 4, are the only ones that show a direct relationship between implicit self-esteem and depression (Phillips et al., 2010). In relation to results on suicide ideation, the sample was insufficiently powered to analyse group differences. However, severity of suicidal ideation was significantly predicted by decreases in implicit self-esteem, above the predictive power of explicit self-esteem, which acted in the same direction (lower explicit self-esteem was related to increased severity of suicidal ideation). This result runs against those obtained by Franck (2007a), who found that only depressed patients without suicidal ideation will show decreased levels of implicit self-esteem, while suicidal patients showed the same level of implicit self-esteem as did non-depressed controls. Franck's design utilised statistical test for comparison of group means, which the study reported in this thesis was not powered to do. However, at a conceptual level, Franck's finding was not replicated. It was necessary to enlarge the sample size to find the relationship between implicit self-esteem and suicide ideation in depressed patients, and reliable results in the relationship between implicit self-esteem and depression.

This was done in the study reported in Chapter 5. The sample for this study was composed of the depressed subgroup of the sample utilised in Chapter 4 and the baseline measures of participants of the REDIT study. Besides attempting to replicate the findings of the previous study, a measure of self-esteem discrepancy was also

scrutinised in its relationship to symptomatic outcomes. This design was powered to 80% to find medium effect sizes at a confidence level of 95% with up to 3 predictors (Cohen, 1992). This study also showed that D scores were not directly related to measures of depression, nor to indices of explicit self-esteem. The previous chapter had found a positive relationship between implicit and explicit self-esteem scores in the prediction of suicide. In this chapter, D scores were unrelated to suicide scales. Analyses showed that most of the prediction of suicidal ideation severity was explained by explicit self-esteem and not by D scores. The study in Chapter 5 counted with greater statistical power, and it is in line with most of the literature finding no direct relationship of D with depression measures including suicide (Phillips et al., 2010). However, more congruently with the existing literature, this study found significant and interesting relationships between symptomatic measures and self-esteem discrepancy, understood as the vectorial difference between explicit and implicit self-esteem. This difference score was related to depressive symptoms and suicidal ideation severity, with a tendency to increase severity of both depression and suicide when implicit self-esteem augments in relation to explicit self-esteem. This result was replicated across different symptomatic measures. Both implicit and explicit self-esteem scores were able to differentiate between suicidal and non-suicidal patients. These results are in line with most of the literature, and represent successful replications indicating that the value of the SE-IAT in the context of the study of depression emerges when implicit and explicit self-esteem are considered in relation to each other, and that both depression and suicide are related to a pattern of discrepant self-esteem, with lower explicit self-esteem and higher implicit self-esteem predicting both severity of depression and suicide (Briñol et al., 2006; Franck et al., 2007a; Pavlickova et al., 2014; Phillips & Hine, 2014; Schröder- Abé et al., 2007a; Valiente et al., 2011). It lends support to the notion that D scores are not directly related to depression (Franck et al., 2007a; Phillips et al., 2010).

Chapter 6 aimed at addressing the same relationships found in the previous one, but in the context of an outcome study. Furthermore, it aimed to ascertain the change in implicit self-esteem associated with treatment and the change in symptomatic outcomes. D scores kept showing robustness to the demographic characteristics of the sample, and allowed a calculation of an index of test-retest reliability that showed a good level and was not related to treatment. D scores were negative for certain groups at different time points (for the control group at baseline and 8 weeks, for the DIT group at 16 weeks), disconfirming the hypothesis of universality of positive implicit self-esteem. However, of all the studies reported in this thesis, this was the only one that did not find positive

D scores for all groups. Following Sterling's logic (1959; Sterling et al., 1995), given that most of the literature informs of positive self-esteem across several domains (Yamaguchi et al., 2007), the result obtained in this chapter could be a chance finding (given that probabilities were set at a 95% confidence level). The reader should remain sceptical about this particular result. It is also important to keep in mind that assuming that a positive *D* score reflecting positive self-esteem implies the assumption that the scores yield by the SE-IAT have a non-arbitrary zero point. In fact, what the SE-IAT measures is the implicit preference for the association between self and good in comparison with the automatic tendency to associate not-me with good (and not-me and bad over me and bad). There is no zero, no psychometric possibility of saying that a person has no implicit self-esteem (Blanton & Jaccard, 2006). Treatment did not achieve significant changes in *D* scores: the SE-IAT remained stable across time in the control group, and across time and in the face of the influence of psychotherapeutic intervention in the treatment group. The stability showed by the SE-IAT was larger than that encountered in the literature, but results in the literature tend to show composite indices of test-retest reliability for a group of thematically distinct IATs (Nosek et al., 2007a). This stability somewhat contradicts studies that show that implicit self-esteem in formerly depressive patients decreases after a negative mood induction (De Raedt et al., 2006; Franck et al., 2008; Gemar et al., 2001). However, there are no studies of therapeutic interventions for depression that make use of implicit measures of self-esteem. Mood manipulations in the literature were carried out immediately before administering the SE-IAT, and the mood manipulation is negative, which is different from therapeutic interventions that search for a positive change in mood. In the case of the study reported in Chapter 6, replications are needed, this being the first study of its type. Results for the effectiveness of the therapeutic intervention are still in preparation, but based on the null change of measures of explicit self-esteem by means of depression treatment, we can surmise that the effectiveness of the treatment was weak in relation to these variables. In any case, this study lends support to the temporal stability of the SE-IAT provided in most of the meta-analytic literature (Bosson et al., 2000; Egloff et al., 2005). Once again, *D* scores were not directly related to depression, also confirming the findings of meta-analytic literature (Buhrmester et al., 2011). More importantly, this chapter showed that implicit and explicit self-esteem are to be used together, either as a discrepancy or an interaction, to successfully differentiate between internalising and externalising problems. Fragile self-esteem has been shown in the literature to be related to externalising reactions, including mania and (non-self-blaming) paranoia (Nakamura

et al., 2015; Pavlickova et al., 2014; Valiente et al., 2011). On the other hand, this chapter, together with the literature, shows that damaged self-esteem is related to internalising manifestations including depression and suicide risk (Cheng et al., 2012; Creemers et al., 2013; Dimaro et al., 2015; Kesting et al., 2011; Pavlickova et al., 2014; Schreiber et al., 2012). It is important to remind the reader here that depression has been found in both fragile and damaged self-esteem configurations (Lemmens et al., 2014; Phillips & Hine, 2014; Vater et al., 2010), which reminds us of the large heterogeneity found in individuals who share this diagnosis (Abramson, Metalsky, & Alloy, 1989; Chen, Eaton, Gallo, & Nestadt, 2000; Weissman et al., 1986). However, suicidal tendencies, which are commonly found in depression, are associated to damaged self-esteem only. Discrepant self-esteem is a risk factor for general psychopathology (Schröder- Abé et al., 2007a; Schröder- Abé et al., 2007b; Wegener et al., 2015), but it is the direction of that discrepancy which seems to distinguish between internalising and externalising manifestations.

Conclusions and Limitations

This thesis is composed of a series of empirical studies using the Implicit Association Test in different samples to demonstrate its versatility and to evaluate its usefulness in different interpersonal domains: attachment transmission and adult depression. Results in the literature appear mixed or incomplete, so this thesis is part of a renewed interest, shared by the scientific community, in developing our knowledge by arriving at scientific statements that will stand up to generalisation and can ultimately become useful in clinical situations outside the laboratory.

With regards to attachment transmission, the whole field of implicit measures in attachment is still very young. The study reported here was the first of its kind, and replication is needed. It was impossible to ascertain relationships between maternal implicit attitudes regarding attachment and parenthood using this sample, and future studies should include measures of parent attachment style to find how implicit attitudes can moderate the relationship between attitudes, parenthood and attachment transmission.

This thesis has found that the SE-IAT adapted to tablet computers is a valid and reliable measure, robust to most demographic variations and with good levels of test-retest and internal reliabilities. According to these psychometric properties, the use of a tablet-based IAT is as reliable as when it is administered on a desktop computer, but

with the added value of portability, therefore widening the spectrum of settings in which this measure can be used, and has the potential to facilitate recruitment for psychological study, sparing participants from the need to travel to a laboratory. The tablet system shows its full potential when taking advantage of mobile communication, allowing data to be collected centrally, anonymously and in real time.

Implicit self-esteem measured with the IAT is a stable construct, related to a steady trait of automatically valuing the self. This trait appears to be unaffected by psychotherapeutic intervention. Regarding the usefulness of the SE-IAT in depression, this thesis can safely conclude that implicit self-esteem scores are not directly related to depressive symptomatology. However, D scores yield valuable information when they are used in conjunction with indicators of explicit self-esteem. The present thesis and the existing literature have shown that the discrepancy between implicit and explicit levels of self-esteem consistently predicts psychopathological outcomes, and that the direction of the discrepancy is useful in distinguishing between externalising and internalising psychopathology. In the particular case of depression, it was found that a tendency to damaged self-esteem predicted severity of depression and suicidal ideation, as well as other internalising problems.

This shows that implicit and explicit self-esteem have a dynamic relationship, which lends support to the conception of dual models of processing self-referential information (Fazio & Olson, 2003), and to the ideas that it is the interplay of levels that ultimately dictates behaviour in different circumstances. Attitudinal models should always allow for the interplay of implicit and explicit aspects, because it has been demonstrated that they are not completely independent from each other, even when they do not correlate, because they do show an interaction when related to symptomatic and behavioural outcomes. However, in order to capture this interplay, measures have to be continuously evaluated; especially implicit measures, whose functioning is not yet completely understood (Karpinski & Steinberg, 2006).

Arriving to the most robust conclusion in this thesis, namely the predictive power of the discrepancy between implicit and explicit self-esteem on severity of depression and suicidal ideation implied trialling the IAT in different context of the socio-cognitive domain. Besides these robust findings, it is important to keep in mind the unfruitful avenues of research also found in through this thesis, particularly the fact that there seems to be no direct (nor moderated) relationship between implicit attitudes towards parenting and attachment, and offspring's attachment status. Also we found no direct

relationship between implicit self-esteem and measures of depression. This thesis was also able to ascertain the methodological integrity of this measure in a heterogeneous adolescent sample, where it showed similar good to excellent psychometric properties as the desktop computer adult versions.

Limitations

Besides the limitations of each particular study, which have been reported above, a limitation transversal to the whole thesis is the lack of a standardised instrument devoted exclusively to the measurement of explicit self-esteem. The studies in this thesis that used explicit self-esteem did it by creating a composite score utilising items from the BDI-II, which are highly correlated with measures of self-esteem in the literature (Heatherton & Polivy, 1991; Osman et al., 1997). This thesis reached the conclusion that implicit self-esteem measured with the IAT is valuable when used together with measures of explicit self-esteem. Future research should always include both measures, in order to assess the interplay between explicit and implicit levels of self-esteem in a standardised and validated manner.

In spite of the results enumerated above, and the fact that the IAT is an increasingly popular instrument, its underlying processes are still unclear and debated (Buhrmester et al., 2011). For the purpose of this thesis and its results, a major limitation is related to the differential nature of the measure itself.

It is assumed that implicit self-esteem is a measure comparable to explicit self-esteem, however that is debatable. It has been found that the IAT can be subsumed under two factors, one with me/positive-other/negative and another with the opposite valence, namely me/negative-other/positive. This is the basic configuration of the compatible and incompatible blocks of the SE-IAT. Obtaining a single score from the difference between these two blocks implies accepting the assumption that compatible and incompatible blocks have the same weight, but the opposite sign when predicting a psychological criterion (Blanton, Jaccard, Gonzales, & Christie, 2006). In a simple linear regression model:

$$Y = \alpha + \beta_{IAT} + \varepsilon$$

where Y is a psychological construct (e.g. depression), α is an intercept and β_{IAT} is the slope coefficient for *D* scores, we could replace the slope term:

$$\beta_{IAT} = \beta_{(compatible - incompatible)}$$

which is the same as presenting the first formula in this way:

$$Y = \alpha + \beta_{compatible} - \beta_{incompatible} + \varepsilon$$

The assumption is then that implicit self-esteem is not merely how rapidly one associates the self with positive traits, but also how rapidly one associates others with negative traits. From this perspective, the SE-IAT reveals as much about participants' self-related attitudes as about participants' attitudes towards the other.

Conceptualising self-esteem in this way might seem counterintuitive at first. However, early conceptualisations of self-esteem included this comparative quality. Leon Festinger's theory of social comparison (1954) has as a main tenet that people construct self-evaluations by comparing themselves with others. This implies that every time a person completes an explicit measure of self-esteem, they are (more or less consciously) thinking of another person (real or fictitious) to compare themselves with. Usually this is a downward comparison to preserve a positive self-evaluation (Wills, 1981). The ability to apprehend ourselves is a developmental achievement that requires the internalisation of the perspective of the other (Fonagy & Target, 1997). From this point of view, we can see self-esteem as related to social affects like guilt, pride or shame. Self-esteem can from this perspective, be conceptualised as a moral orientation towards oneself, rooted in (at least partially) shared ideals of what is good and worthy (Tafarodi, 2006).

Even considering these theories, the question remains: is higher implicit self-esteem a positive regard to oneself or a negative regard to the other? Some of the authors who developed the IAT tend to think that high implicit self-esteem is both (Greenwald & Banaji, 1995), while others have tackled the problem and developed the GNAT, as it was reviewed in Chapter 1 (Nosek & Banaji, 2001). In the particular case of depression with high levels of implicit self-esteem, we can surmise that the depressed individual is very quickly associating not-me with negative attributes or himself with positive ones. This last factor is the assumption for the test, and the one that gives name to it: implicit self-esteem openly suggests a self-evaluation, and not the evaluation of the other. Future research with this measure should utilise, besides the accepted scoring algorithm, the analysis plan used by De Raedt (De Raedt et al., 2006), which attempts at comparing positive and negative trials within the domains of "me" and "not-me"

separately. This should be used in conjunction with the accepted scoring algorithm, in order to compare the results and to ascertain which method is superior. It will be very interesting to see if increased implicit self-esteem in depressed patients correspond to a positive automatic bias towards the self, or the automatic criticism of an internalised "not-me", which would correspond to the classical definition of depressive disorders (Freud, 1917).

Lastly, the way that discrepancy scores have been calculated might cause confusion because the result does not yield information about the absolute levels of explicit and implicit self-esteem but only their relative difference. This means that, in a hypothetic case, an individual who shows a score of zero in explicit self esteem (consider this score an indicator of low self-esteem for this example) and an implicit self-esteem score of -1 (consider this an indicator of low implicit self-esteem) will have the same discrepancy score of an individual who scores 2 in explicit self-esteem (a value that for this example represents a very high self-esteem) and 1 in implicit self-esteem (a high level for this example), and to an individual who scores -1 in explicit self-esteem (low) and -2 in implicit self-esteem (very low). Are these individuals the same when it comes to their levels of self-esteem? It would be irresponsible to answer that question affirmatively. These hypothetical individuals are qualitatively distinct. However, they show a quantitatively comparable risk for depressive symptoms, suicide, and other internalising problems. This is because the measure of discrepancy is necessarily within subjects, and it is the discrepancy itself that signals vulnerability. In depression particularly, but also in other disorders, one depressed patient is different to the next depressed patient. More than a limitation or a contradiction, the results obtained using this discrepancy measure is an illustration of the nature of our scientific object, the human mind, and its immense diversity

References

Aberson, & Beeney. (2007). Does Substance Use Affect Reliabilities of the Implicit Association Test? *The Journal of Social Psychology, 147*(1), 27-40. doi: 10.3200/socp.147.1.27-40

Abramson, Metalsky, & Alloy. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review, 96*(2), 358.

Adam, Gunnar, & Tanaka. (2004). Adult Attachment, Parent Emotion, and Observed Parenting Behavior: Mediator and Moderator Models. *Child Development, 75*(1), 110-122. doi: 10.2307/3696569

Aidman, & Carroll. (2003). Implicit individual differences: relationships between implicit self‐esteem, gender identity, and gender attitudes. *European Journal of Personality, 17*(1), 19-37.

Ainsworth, Blehar, Waters, & Wall. (1978). *Patterns of attachment: A psychological study of the strange situation*. Oxford, England: Lawrence Erlbaum.

Allport. (1935). Attitudes. In C. A. Murchison (Ed.), *Handbook of social psychology* (pp. 798-844). Worcester, MA: Clark University Press.

Altmann. (2008). *Attitude: a concept analysis.* Paper presented at the Nursing Forum.

American Psychiatric Association. (2010). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association [APA]. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Arkes, & Tetlock. (2004). Attributions of implicit prejudice, or" would Jesse Jackson'fail'the Implicit Association Test?". *Psychological Inquiry, 15*(4), 257-278.

Arnone, McIntosh, Ebmeier, Munafo, & Anderson. (2012). Magnetic resonance imaging studies in unipolar depression: systematic review and meta-regression analyses. *European Neuropsychopharmacology, 22*(1), 1-16. doi: 10.1016/j.euroneuro.2011.05.003

Ashburn-Nardo, Voils, & Monteith. (2001). Implicit associations as the seeds of intergroup bias: How easily do they take root? *Journal of Personality and Social Psychology, 81*(5), 789.

Baccus, Baldwin, & Packer. (2004). Increasing implicit self-esteem through classical conditioning. *Psychological Science, 15*(7), 498-502. doi: Doi 10.1111/J.0956-7976.2004.00708.X

Back, Schmukle, & Egloff. (2009). Predicting actual behavior from the explicit and implicit self-concept of personality. *Journal of Personality and Social Psychology, 97*(3), 533. doi: 10.1037/a0016229

Bagby, Ryder, Schuller, & Marshall. (2014). The Hamilton Depression Rating Scale: has the gold standard become a lead weight?

Baldwin. (1992). Relational Schemas and the Processing of Social Information. *Psychological Bulletin, 112*(3), 461-484. doi: Doi 10.1037/0033-2909.112.3.461

Banaji, Hardin, & Rothman. (1993). Implicit stereotyping in person judgment. *Journal of Personality and Social Psychology, 65*(2), 272. doi: Doi 10.1037/0022-3514.65.2.272

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

Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & Van Ijzendoorn. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychological Bulletin, 133*(1), 1. doi: 10.1037/0033-2909.133.1.1

Bargh, Chaiken, Govender, & Pratto. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology, 62*(6), 893.

Barnes-Holmes, Murtagh, Barnes-Holmes, & Stewart. (2011). Using the Implicit Association Test and the Implicit Relational Assessment Procedure to measure attitudes toward meat and vegetables in vegetarians and meat-eaters. *The Psychological Record, 60*(2), 6.

Baron, & Banaji. (2006). The development of implicit attitudes evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological Science, 17*(1), 53-58. doi: 10.1111/j.1467-9280.2005.01664.x

Bartholomew, & Horowitz. (1991). Attachment Styles Among Young Adults: A Test of a Four-Category Model. *Journal of Personality & Social Psychology, 61*(2), 226-244.

Bartholomew, & Moretti. (2002). The dynamics of measuring attachment. *Attachment & Human Development, 4*(2), 162-165. doi: 10.1080/14616730210157493

Bartling, Wiefel, Klapp, Jonkman, & Lenz. (2006). *Mother-child interaction and child's neuroendocrine stressregulation*. Paper presented at the World Association for Infant Mental Health Conference, Paris. Thesis.docx

Bartz, & Lydon. (2004). Close Relationships and the Working Self-Concept: Implicit and Explicit Effects of Priming Attachment on Agency and Communion. *Personality and Social Psychology Bulletin, 30*(11), 1389-1401. doi: 10.1177/0146167204264245

Baumeister. (1990). Suicide as escape from self. *Psychological Review, 97*(1), 90. doi: Doi 10.1037//0033-295x.97.1.90

Baumeister, Campbell, Krueger, & Vohs. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological science in the public interest, 4*(1), 1-44.

Baumeister, Smart, & Boden. (1996). Relation of threatened egotism to violence and aggression: the dark side of high self-esteem. *Psychological Review, 103*(1), 5.

Beck. (2008). The evolution of the cognitive model of depression and its neurobiological correlates. *Am J Psychiatry, 165*(8), 969-977. doi: 10.1176/appi.ajp.2008.08050721

Beck, & Alford. (1967). *Depression: Causes and treatment*: University of Pennsylvania Press.

Beck, Steer, & Brown. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.

Bednar, & Peterson. (1995). Self-Esteem: Paradoxes and Innovations in Clinical Theory and Practice, American Psychological Association. *Washington, DC*.

Belmaker, & Agam. (2008). Major depressive disorder. *New England Journal of Medicine, 358*(1), 55-68. doi: 10.1056/NEJMra073096

Belsky. (2005). The developmental and evolutionary psychology of intergenerational transmission of attachment. *Attachment and bonding: A new synthesis*, 169-198.

Belsky, Rovine, & Taylor. (1984). The Pennsylvania Infant and Family Development Project, III: The origins of individual differences in infant-mother attachment: Maternal and infant contributions. *Child Development, 55*(3), 718-728.

Berthelot, Ensink, Bernazzani, Normandin, Fonagy, & Luyten. (2014). Intergenerational Transmission of Attachment in Abused and Neglected Mothers: The Role of Trauma-Specific Reflective Functioning. *Infant Mental Health Journal*.

Blair. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6*(3), 242-261. doi: Doi 10.1207/S15327957pspr0603_8

Blair, Ma, & Lenton. (2001). Imagining stereotypes away: the moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology, 81*(5), 828. doi: 10.1037//0022-3514.81.5.828

Blair-West, Cantor, Mellsop, & Eyeson-Annan. (1999). Lifetime suicide risk in major depression: sex and age determinants. *Journal of Affective Disorders, 55*(2), 171-178.

Blair‐West, Mellsop, & Eyeson‐Annan. (1997). Down‐rating lifetime suicide risk in major depression. *Acta Psychiatrica Scandinavica, 95*(3), 259-263.

Blanton, & Jaccard. (2006). Arbitrary metrics in psychology. *American Psychologist, 61*(1), 27.

Blanton, Jaccard, Gonzales, & Christie. (2006). Decoding the implicit association test: Implications for criterion prediction. *Journal of Experimental Social Psychology, 42*(2), 192-212.

Blascovich, & Tomaka. (1991). Measures of self-esteem. In J. P. Robinson & P. R. Shaver (Eds.), *Measures of personality and social psychological attitudes* (Vol. 1, pp. 115-160). San Diego, CA: Academic Press.

Blass. (2010). Affirming 'That's not psycho‐analysis!' On the value of the politically incorrect act of attempting to define the limits of our field. *The International Journal of Psychoanalysis, 91*(1), 81-89.

Blass, & Carmeli. (2007). The case against neuropsychoanalysis: On fallacies underlying psychoanalysis' latest scientifi c trend and its negative impact on psychoanalytic discourse. *The International Journal of Psychoanalysis, 88*(1), 19-40.

Blatt. (2011). Two configurations of personality development and psychopathology: Etiologic and therapeutic implications. In C. Piers (Ed.), *Personality and Psychopathology*. New York, Dordrecht, Heildelberg, London: Springer.

Block, & Robins. (1993). A Longitudinal Study of Consistency and Change in Self-Esteem from Early Adolescence to Early Adulthood. *Child Development, 64*(3), 909-923. doi: 10.2307/1131226

Bolognini, Plancherel, Bettschart, & Halfon. (1996). Self-esteem and mental health in early adolescence: Development and gender differences. *Journal of adolescence, 19*(3), 233-245.

Boom. (1994). The influence of temperament and mothering on attachment and exploration: An experimental manipulation of sensitive responsiveness among lower‑class mothers with irritable infants. *Child Development, 65*(5), 1457-1477.

Bornstein, & D'Agostino. (1992). Stimulus recognition and the mere exposure effect. *Journal of Personality and Social Psychology, 63*(4), 545.

Bos, Huijding, Muris, Vogel, & Biesheuvel. (2010). Global, contingent and implicit self-esteem and psychopathological symptoms in adolescents. *Personality and Individual Differences, 48*(3), 311-316.

Bosson. (2006). Assessing Self-Esteem via Self- Reports and Nonreactive Instruments: Issues and Recommendations. In M. H. Kernis (Ed.), *Self-esteem issues and answers: A sourcebook of current perspectives*. Hove: Psychology Press.

Bosson, Brown, Zeigler-Hill, & Swann. (2003). Self-enhancement tendencies among people with high explicit self-esteem: The moderating role of implicit self-esteem. *Self and Identity, 2*(3), 169-187.

Bosson, Lakey, Campbell, Zeigler‑Hill, Jordan, & Kernis. (2008). Untangling the Links between Narcissism and Self‑esteem: A Theoretical and Empirical Review. *Social and Personality Psychology Compass, 2*(3), 1415-1439.

Bosson, Swann, & Pennebaker. (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.

Boulet, & Boss. (1991). Reliability and Validity of the Brief Symptom Inventory. *Psychological Assessment, 3*(3), 433-437.

Bowlby. (1969). *Attachment and Loss. Volumen 1: Attachment* (Vol. 2: Attachment). London: The Hogarth Press and the Institute of Psychoanalysis.

Bowlby. (1973). *Attachment and Loss. Volume 2: Separation-anxiety and anger* (Vol. 2: Separation-anxiety and anger). New York: Basic Books.

Brandtstädter, & Rothermund. (2002). The life-course dynamics of goal pursuit and goal adjustment: A two-process framework. *Developmental Review, 22*(1), 117-150.

Brauer, de Jong, Huijding, Laan, & Ter Kuile. (2009). Automatic and deliberate affective associations with sexual stimuli in women with superficial dyspareunia. *Archives of sexual behavior, 38*(4), 486-497.

Brennan, Clark, & Shaver. (1998). Self-report measurement of adult attachment: An integrative overview. *Attachment theory and close relationships*, 46-76.

Breuer, & Freud. (1895). *Studies on Hysteria. The Standard Edition of the Complete Psychological Works of Sigmund Freud, Vol. 2*. London: Hogarth Press and the Institute of Psychoanalysis.

Brewer. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin, 86*(2), 307.

Briñol, Petty, & Wheeler. (2006). Discrepancies between explicit and implicit self-concepts: Consequences for information processing. *Journal of Personality and Social Psychology, 91*(1), 154.

Brown. (2011). *A sociological analysis of maternal ambivalence: class and race differences among new mothers (Unpublished Doctoral Dissertation).* (Doctor of Philosophy), The State University of New Jersey, New Jersey.

Brown, & Bosson. (2001). Narcissus meets Sisyphus: Self-love, self-loathing, and the never-ending pursuit of self-worth. *Psychological Inquiry*, 210-213.

Brown, Mangelsdorf, & Neff. (2012). Father involvement, paternal sensitivity, and father− child attachment security in the first 3 years. *Journal of Family Psychology, 26*(3), 421.

Bruder. (1983). Cerebral laterality and psychopathology: a review of dichotic listening studies. *Schizophrenia Bulletin, 9*(1), 134.

Brunel, Tietje, & Greenwald. (2004). Is the implicit association test a valid and valuable measure of implicit consumer social cognition? *Marketing*, 4.

Budd. (2001). "No sex please-we're British": Sexuality in English and French psychoanalysis. In C. Harding (Ed.), *Sexuality: Psychoanalytic Perspectives* (pp. 52-68). New York: Brunner-Routledge.

Buhlmann, Teachman, Gerbershagen, Kikul, & Rief. (2008). Implicit and explicit self-esteem and attractiveness beliefs among individuals with body dysmorphic disorder. *Cognitive Therapy and Research, 32*(2), 213-225.

Buhrmester, Blanton, & Swann. (2011). Implicit Self-Esteem: Nature, Measurement, and a New Way Forward. *Journal of Personality & Social Psychology, 100*(2), 365-385.

Busch, Fawcett, & Jacobs. (2003). Clinical correlates of inpatient suicide. *The Journal of clinical psychiatry*(64), 14-19.

Byrne. (1961). Interpersonal attraction and attitude similarity. *The Journal of Abnormal and Social Psychology, 62*(3), 713.

Campbell. (1950). The indirect assessment of social attitudes. *Psychological Bulletin, 47*(1), 15.

Campbell, Bosson, Goheen, Lakey, & Kernis. (2007). Do narcissists dislike themselves "deep down inside"? *Psychological Science, 18*(3), 227-229.

Campbell, Goldstein, Schaefer, & Ramey. (1991). Parental beliefs and values related to family risk, educational intervention, and child academic competence. *Early Childhood Research Quarterly, 6*(2), 167-182. doi: http://dx.doi.org/10.1016/0885-2006(91)90005-6

Caspi, Houts, Belsky, Goldman-Mellor, Harrington, Israel, . . . Poulton. (2014). The p factor one general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science, 2*(2), 119-137.

Cesario. (2014a). *Menstrual cycle effect on outgroup male bias*. Unpublished. Retrieved from http://www.PsychFileDrawer.org/replication.php?attempt=MjAx

Cesario. (2014b). Priming, replication, and the hardest science. *Perspectives on Psychological Science, 9*(1), 40-48.

Chee, Sriram, Soon, & Lee. (2000). Dorsolateral prefrontal cortex and the implicit association of concepts and attributes. *Neuroreport, 11*(1), 135-140.

Chen, Eaton, Gallo, & Nestadt. (2000). Understanding the heterogeneity of depression through the triad of symptoms, course and risk factors: a longitudinal, population-based study. *Journal of Affective Disorders, 59*(1), 1-11.

Cheng, & Furnham. (2003). Personality, self-esteem, and demographic predictions of happiness and depression. *Personality and Individual Differences, 34*(6), 921-942.

Cheng, Govorun, & Chartrand. (2012). Effect of self-awareness on negative affect among individuals with discrepant low self-esteem. *Self and Identity, 11*(3), 304-316.

Clark, Teasdale, Broadbent, & Martin. (1983). Effect of mood on lexical decisions. *Bulletin of the Psychonomic Society, 21*(3), 175-178.

Clerkin, & Teachman. (2009). Automatic and strategic measures as predictors of mirror gazing among individuals with body dysmorphic disorder symptoms. *The Journal of Nervous and Mental Disease, 197*(8), 589.

Clerkin, & Teachman. (2010). Training implicit social anxiety associations: An experimental intervention. *Journal of Anxiety Disorders, 24*(3), 300-308. doi: 10.1016/j.janxdis.2010.01.001

Cockerham, Stopa, Bell, & Gregg. (2009). Implicit self-esteem in bulimia nervosa. *Journal of Behavior Therapy and Experimental Psychiatry, 40*(2), 265-273.

Cohen. (1992). A power primer. *Psychological Bulletin, 112*(1), 155.

Conner, & Barrett. (2005). Implicit self-attitudes predict spontaneous affect in daily life. *Emotion, 5*(4), 476.

Covington. (2014). *Implicit Associations in Blood-Injection-Injury Phobia: Changes at Follow-up to Exposure.* (Master of Arts in Psychology), American University, WASHINGTON D.C. Retrieved from http://aladinrc.wrlc.org/bitstream/handle/1961/15275/Covington_american_0008 N_10523display.pdf?sequence=1

Craeynest, Crombez, De Houwer, Deforche, & De Bourdeaudhuij. (2006). Do children with obesity implicitly identify with sedentariness and fat food? *Journal of Behavior Therapy and Experimental Psychiatry, 37*(4), 347-357.

Craeynest, Crombez, Deforche, Tanghe, & De Bourdeaudhuij. (2008). The role of implicit attitudes towards food and physical activity in the treatment of youth obesity. *Eating Behaviors, 9*(1), 41-51.

Craeynest, Crombez, Haerens, & De Bourdeaudhuij. (2007). Do overweight youngsters like food more than lean peers? Assessing their implicit attitudes with a personalized Implicit Association Task. *Food Quality and Preference, 18*(8), 1077-1084.

Creemers. (2014). *Implicit and explicit self-attitudes in relation to adolescent and young adult depression, stress and treatment*: [Sl: sn].

Creemers, Scholte, Engels, Prinstein, & Wiers. (2012). Implicit and explicit self-esteem as concurrent predictors of suicidal ideation, depressive symptoms, and loneliness. *J Behav Ther Exp Psychiatry, 43*(1), 638-646. doi: S0005-7916(11)00093-0 [pii]

10.1016/j.jbtep.2011.09.006

Creemers, Scholte, Engels, Prinstein, & Wiers. (2013). Damaged self-esteem is associated with internalizing problems. *Frontiers in psychology, 4.*

Crocker, Thompson, McGraw, & Ingerman. (1987). Downward comparison, prejudice, and evaluations of others: effects of self-esteem and threat. *Journal of Personality and Social Psychology, 52*(5), 907.

Cunningham, Johnson, Raye, Gatenby, Gore, & Banaji. (2004). Separable neural components in the processing of black and white faces. *Psychological Science, 15*(12), 806-813.

Cunningham, Preacher, & Banaji. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. *Psychological Science, 12*(2), 163-170.

Cusin, Yang, Yeung, & Fava. (2010). Rating scales for depression. In L. Baer & M. A. Blais (Eds.), *Handbook of clinical rating scales and assessment in psychiatry and mental health* (pp. 7-35): Springer.

Cyr, Dubois-Comtois, Pascuzzo, Béliveau, & Ellen. (2014). Transmission of Attachment at Preschool-Age: The Mediating Role of Mother-Child Conversation Styles. *J Child Adolesc Behav, 5*(171), 2.

Dasgupta, & Greenwald. (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.

Dasgupta, McGhee, Greenwald, & Banaji. (2000). Automatic preference for White Americans: Eliminating the familiarity explanation. *Journal of Experimental Social Psychology, 36*(3), 316-328.

De Houwer. (2002). The Implicit Association Test as a tool for studying dysfunctional associations in psychopathology: strengths and limitations. *Journal of Behavior Therapy and Experimental Psychiatry, 33*, 115-133.

De Houwer. (2006). What are implicit measures and why are we using them. In R. W. Wiers & A. W. Stacy (Eds.), *The handbook of implicit cognition and addiction* (pp. 11-28). Thousand Oaks, CA: Sage Publishers.

De Houwer Dirk Hermans Paul Eelen. (1998). Affective and identity priming with episodically associated stimuli. *Cognition & Emotion, 12*(2), 145-169.

De Houwer, & Eelen. (1998). An affective variant of the Simon paradigm. *Cognition & Emotion, 12*(1), 45-62.

De Houwer, Teige-Mocigemba, Spruyt, & Moors. (2009). Implicit Measures: A Normative Analysis and Review. *Psychological Bulletin, 135*(3), 347-368.

de Jong. (2002). Implicit self-esteem and social anxiety: Differential self-favouring effects in high and low anxious individuals. *Behaviour Research and Therapy, 40*(5), 501-508.

De Jong, Pasman, Kindt, & Van den Hout. (2001). A reaction time paradigm to assess (implicit) complaint-specific dysfunctional beliefs. *Behaviour Research and Therapy, 39*(1), 101-113.

De Jong, Sportel, De Hullu, & Nauta. (2012). Co-occurrence of social anxiety and depression symptoms in adolescence: differential links with implicit and explicit self-esteem? *Psychological Medicine, 42*(03), 475-484.

de Jong, van den Hout, Rietbroek, & Huijding. (2003). Dissociations between implicit and explicit attitudes toward phobic stimuli. *Cognition & Emotion, 17*(4), 521-545.

De Man, & Gutierrez. (2002). The relationship between level of self-esteem and suicidal ideation with stability of self-esteem as moderator. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 34*(4), 235.

De Raedt, Franck, Fannes, & Verstraeten. (2008). Is the relationship between frontal EEG alpha asymmetry and depression mediated by implicit or explicit self-esteem? *Biological Psychology, 77*(1), 89-92.

De Raedt, Schacht, Franck, & De Houwer. (2006). Self-esteem and depression revisited: Implicit positive self-esteem in depressed patients? *Behaviour Research and Therapy, 44*(7), 1017-1028. doi: 10.1016/j.brat.2005.08.003

DeHart, & Pelham. (2007). Fluctuations in state implicit self-esteem in response to daily negative events. *Journal of Experimental Social Psychology, 43*(1), 157-165.

DeHart, Pelham, & Tennen. (2006). What lies beneath: Parenting style and implicit self-esteem. *Journal of Experimental Social Psychology, 42*(1), 1-17.

DeHart, Peña, & Tennen. (2013). The development of explicit and implicit self-esteem and their role in psychological adjustment. *Self-esteem: current issues in social psychology, Psychology Press, New York*, 99-123.

Del Carmen, Pedersen, Huffman, & Bryan. (1993). Dyadic distress management predicts subsequent security of attachment. *Infant Behavior and Development, 16*(2), 131-147.

DeMeis, Hock, & McBride. (1986). The Balance of Employment and Motherhood: Longitudinal Study of Mothers' Feelings About Separation From Their First-Born Infants. *Developmental Psychology, 22*(5), 627-632.

Dentale, San Martini, De Coro, & Di Pomponio. (2010). Alexithymia increases the discordance between implicit and explicit self-esteem. *Personality and Individual Differences, 49*(7), 762-767.

Dentale, Vecchione, Coro, & Barbaranelli. (2012). On the relationship between implicit and explicit self-esteem: The moderating role of dismissing attachment. *Personality and Individual Differences, 52*(2), 173-177.

Derogatis. (1992). *SCL-90-R: Administration, Scoring of Procedures Manual-II for the R (evised) Version and Other Instruments of the Psychopathology Rating Scale Series*: Clinical Psychometric Research Incorporated.

Derogatis. (1993). *BSI, brief symptom inventory: Administration, scoring & procedures manual*: National Computer Systems.

Derogatis, & Melisaratos. (1983). The Brief Symptom Inventory: An Introductory Report. *Psychological Medicine, 13*, 595-605.

Dewitte, De Houwer, & Buysse. (2008). On the role of the implicit self-concept in adult attachment. *European Journal of Psychological Assessment, Vol.24*(4), pp.

Diener, Kuehner, Brusniak, Ubl, Wessa, & Flor. (2012). A meta-analysis of neurofunctional imaging studies of emotion and cognition in major depression. *Neuroimage, 61*(3), 677-685.

Diesing. (1985). Comments on "Why Freud's research methodology was unscientific" by von Eckhardt. *Psychoanalysis & Contemporary Thought, 8*(4), 551-566.

Dijksterhuis. (2004). I like myself but I don't know why: enhancing implicit self-esteem by subliminal evaluative conditioning. *Journal of Personality and Social Psychology, 86*(2), 345.

Dijksterhuis, & Van Knippenberg. (1998). The relation between perception and behavior, or how to win a game of trivial pursuit. *Journal of Personality and Social Psychology, 74*(4), 865.

Dimaro, Roberts, Moghaddam, Dawson, Brown, & Reuber. (2015). Implicit and explicit self-esteem discrepancies in people with psychogenic nonepileptic seizures. *Epilepsy & Behavior.*

Donges, Jachmann, Kersting, Egloff, & Suslow. (2015). Attachment anxiety and implicit self-concept of neuroticism: Associations in women but not men. *Personality and Individual Differences, 72*, 208-213.

Donnellan, Trzesniewski, Robins, Moffitt, & Caspi. (2005). Low self-esteem is related to aggression, antisocial behavior, and delinquency. *Psychological Science, 16*(4), 328-335.

Dozois, Dobson, & Ahnberg. (1998). A psychometric evaluation of the Beck Depression Inventory–II. *Psychological Assessment, 10*(2), 83.

Draine, & Greenwald. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General, 127*(3), 286.

DuBois, Felner, Brand, & George. (1999). Profiles of self-esteem in early adolescence: Identification and investigation of adaptive correlates. *American Journal of Community Psychology, 27*(6), 899-932.

Dubois, & Tevendale. (1999). Self-esteem in childhood and adolescence: Vaccine or epiphenomenon? *Applied and Preventive Psychology, 8*(2), 103-117.

Dykas, & Cassidy. (2011). Attachment and the Processing of Social Information Across the Life Span: Theory and Evidence. *Psychological Bulletin, 137*(1), 19-46.

Early Child Care Research Network NICHD. (1994). Child care and child development: the NICHD study of early child care. In S. L. Friedman & H. C. Haywood (Eds.), *Developmental Follow-Up: Concepts, Domains and Methods* (pp. 377-396). New York: Academic Press.

Easterbrook, Gopalan, Berlin, & Matthews. (1991). Publication bias in clinical research. *The Lancet, 337*(8746), 867-872.

Eaton, Shao, Nestadt, Lee, Bienvenu, & Zandi. (2008). Population-Based Study of First Onset and Chronicity in Major Depressive Disorder. *Archives of General Psychiatry, 65*(5), 513-520. doi: 10.1001/archpsyc.65.5.513

Egenolf, Stein, Koenig, Holtforth, Dierks, & Caspar. (2013). Tracking the implicit self using event-related potentials. *Cognitive, Affective, & Behavioral Neuroscience, 13*(4), 885-899.

Egloff, & Schmukle. (2002). Predictive validity of an implicit association test for assessing anxiety. *Journal of Personality and Social Psychology, 83*(6), 1441.

Egloff, & Schmukle. (2003). Does social desirability moderate the relationship between implicit and explicit anxiety measures? *Personality and Individual Differences, 35*(7), 1697-1706.

Egloff, Schwerdtfeger, & Schmukle. (2005). Temporal stability of the implicit association test-anxiety. *Journal of Personality Assessment, 84*(1), 82-88.

Elgersma, Glashouwer, Bockting, Penninx, & de Jong. (2013). Hidden scars in depression? Implicit and explicit self-associations following recurrent depressive episodes. *Journal of Abnormal Psychology, 122*(4), 951.

Ellwart, Rinck, & Becker. (2006). From fear to love: individual differences in implicit spider associations. *Emotion, 6*(1), 18.

Emmons. (1987). Narcissism: theory and measurement. *Journal of Personality and Social Psychology, 52*(1), 11.

Engelhard, Huijding, van den Hout, & de Jong. (2007). Vulnerability associations and symptoms of post-traumatic stress disorder in soldiers deployed to Iraq. *Behaviour Research and Therapy, 45*(10), 2317-2325.

Epstein. (1994). Integration of the Cognitive and the Psychodynamic Unconscious. *American Psychologist, 49*(8), 709-724.

Epstein, & Morling. (1995). Is the self motivated to do more than enhance and/or verify itself? *Efficacy, agency, and self-esteem* (pp. 9-29): Springer.

Erol, & Orth. (2011). Self-esteem development from age 14 to 30 years: a longitudinal study. *Journal of Personality and Social Psychology, 101*(3), 607.

Everyday Health. (2013). Stats and facts about depression in America. from http://www.everydayhealth.com/health-report/majordepression/depression-statistics.aspx

Fanelli. (2010). Positive results increase down the Hierarchy of the Sciences. *PLoS One, 5*(3).

Fanelli. (2011). Negative results are disappearing from most disciplines and countries. *Scientometrics, 90*(3), 891-904.

Fazio. (1986). How do attitudes guide behavior. *Handbook of motivation and cognition: Foundations of social behavior, 1*, 204-243.

Fazio. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. *Advances in experimental social psychology, 23*(75-109).

Fazio, Jackson, Dunton, & Williams. (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.

Fazio, & Olson. (2003). Implicit Measures in Social Cognition Research: Their Meaning and Use. *Annual Review of Psychology, 54*(1), 297-327. doi: doi:10.1146/annurev.psych.54.101601.145225

Fazio, Powell, & Williams. (1989). The role of attitude accessibility in the attitude-to-behavior process. *Journal of consumer research*, 280-288.

Fearon, Shmueli‐Goetz, Viding, Fonagy, & Plomin. (2014). Genetic and environmental influences on adolescent attachment. *Journal of Child Psychology and Psychiatry, 55*(9), 1033-1041.

Feldman. (1998). *Coding interactive behavior manual*. Israel: Bar-Ilan University.

Feldman. (2000). Parents' convergence on sharing and marital satisfaction, father involvement, and parent–child relationship at the transition to parenthood. *Infant Mental Health Journal, 21*(3), 176-191. doi: 10.1002/1097-0355(200007)21:3<176::aid-imhj3>3.0.co;2-4

Feldman, Greenbaum, Mayes, & Erlich. (1997). Change in mother-infant interactive behavior: Relations to change in the mother, the infant, and the social context. *Infant Behavior and Development, 20*(2), 151-163. doi: http://dx.doi.org/10.1016/S0163-6383(97)90018-7

Feldman, Keren, Gross-Rozval, & Tyano. (2004). Mother-Child Touch Patterns in Infant Feeding Disorders: Relation to Maternal, Child, and Environmental Factors. *Journal of the American Academy of Child & Adolescent Psychiatry, 43*(9), 1089-1097. doi: http://dx.doi.org/10.1097/01.chi.0000132810.98922.83

Feldman, Weller, Sirota, & Eidelman. (2003). Testing a family intervention hypothesis: the contribution of mother-infant skin-to-skin contact (kangaroo care) to family interaction, proximity, and touch. *Journal of Family Psychology, 17*(1), 94.

Ferber, & Feldman. (2005). Delivery Pain and the Development of Mother—Infant Interaction. *Infancy, 8*(1), 43-62.

Ferguson. (1935). The influence of individual attitudes on construction of an attitude scale. *The Journal of Social Psychology, 6*(1), 115-117.

Festinger. (1954). A theory of social comparison processes. *Human relations, 7*(2), 117-140.

Fitzgerald, Laird, Maller, & Daskalakis. (2008). A meta‐analytic study of changes in brain activation in depression. *Human brain mapping, 29*(6), 683-695.

Fleming, & Courtney. (1984). The dimensionality of self-esteem: II. Hierarchical facet model for revised measurement scales. *Journal of Personality and Social Psychology, 46*(2), 404.

Fonagy. (1982). The integration of psychoanalysis and experimental science: A review. *International Review of Psycho-Analysis*.

Fonagy, Gergely, Target & Jurist. (2004). *Affect regulation, mentalization and the development of the self*: Karnac books.

Fonagy, Luyten, & Perkins. (2015). Development and Validation of a Self-report Measure of Mentalizing: The Reflective Functioning Questionnaire. *Unpublished manuscript*.

Fonagy, Steele, & Steele. (1991). Maternal Representations of Attachment during Pregnancy Predict the Organization of Infant-Mother Attachment at One Year of Age. *Child Development, 62*(5), 891-905. doi: 10.2307/1131141

Fonagy, & Target. (1997). Attachment and reflective function: Their role in self-organization. *Development and Psychopathology, 9*(04), 679-700.

Fox, Mathews, Calder, & Yiend. (2007). Anxiety and sensitivity to gaze direction in emotionally expressive faces. *Emotion, 7*(3), 478.

Fraley, Waller, & Brennan. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology, 78*(2), 350.

Franck, De Raedt, & De Houwer. (2008). Activation of latent self-schemas as a cognitive vulnerability factor for depression: The potential role of implicit self-esteem. *Cognition & Emotion, 22*(8), 1588-1599. doi: 10.1080/02699930801921271

Franck, De Raedt, Dereu, & Van den Abbeele. (2007a). Implicit and explicit self-esteem in currently depressed individuals with and without suicidal ideation. *Journal of Behavior Therapy and Experimental Psychiatry, 38*(1), 75-85. doi: 10.1016/j.jbtep.2006.05.003

Franck, Raedt, & Houwer. (2007b). Implicit but not explicit self-esteem predicts future depressive symptomatology. *Behaviour Research and Therapy, 45*(10), 2448-2455.

Freud. (1896). Further remarks on the neuro-psychoses of defence. In J. Stratchey (Ed.), *The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume III (1893-1899): Early Psycho-Analytic Publications* (Vol. III, pp. 157-185). London: The Hogarth Press and the Institue of Psychoanalysis.

Freud. (1914). *On narcissism: an introduction. The Standard Edition of the Complete Psychological Works of Sigmund Freud, Vol. 10*. London: Hogarth Press and the Institute of Psychoanalysis.

Freud. (1915). The unconscious *The standard edition of the complete psychological works of Sigmund Freud, Volume XIV (1914-1916): On the history of the psycho-analytic movement, paper on metapsychology and other works* (Vol. The Hogarth Press and the Institute of Psychoanalysis): London.

Freud. (1916). *Introductory lectures on psychoanalysis. The Standard Edition of the Complete Psychological Works of Sigmund Freud, Vol. 16*. London: Hogarth Press and the Institute of Psychoanalysis.

Freud. (1917). Mourning and Melancholia *The Standard Edition of the Complete Psychological Works of Sigmund Freud,*
Volume XIV (1914-1916): On the History of the Psycho-Analytic Movement, Papers on Metapsychology and Other Works (pp. 237-258). London: The Hogarth Press and The Institue of Psychoanalysis.

Friese, Hofmann, & Schmitt. (2009). When and why do implicit measures predict behaviour? Empirical evidence for the moderating role of opportunity, motivation, and process reliance. *European Review of Social Psychology, 19*(1), 285-338. doi: 10.1080/10463280802556958

Furnham, Badmin, & Sneade. (2002). Body image dissatisfaction: Gender differences in eating attitudes, self-esteem, and reasons for exercise. *The Journal of psychology, 136*(6), 581-596.

Gamer, Schmukle, Luka-Krausgrill, & Egloff. (2008). Examining the Dynamics of the Implicit and the Explicit Self-Concept in Social Anxiety: Changes in the Implicit Association Test–Anxiety and the Social Phobia Anxiety Inventory Following Treatment. *Journal of Personality Assessment, 90*(5), 476-480. doi: 10.1080/00223890802248786

Gawronski. (2002). What does the Implicit Association Test measure? A test of the convergent and discriminant validity of prejudice-related IATs. *Experimental Psychology, 49*(3), 171-180.

Gawronski. (2007). Editorial: Attitudes can be measured! But what is an attitude? *Social Cognition, 25*(5), 573-581.

Gawronski. (2009). Ten frequently asked questions about implicit measures and their frequently supposed, but not entirely correct answers. *Canadian Psychology/Psychologie canadienne, 50*(3), 141.

Gawronski, Hofmann, & Wilbur. (2006). Are "implicit" attitudes unconscious? *Consciousness and Cognition, 15*(3), 485-499. doi: http://dx.doi.org/10.1016/j.concog.2005.11.007

Geller, Johnston, Madsen, Goldner, Remick, & Birmingham. (1998). Shape‐and weight‐based self‐esteem and the eating disorders. *International Journal of Eating Disorders, 24*(3), 285-298.

Gemar, Segal, Sagrati, & Kennedy. (2001). Mood-induced changes on the Implicit Association Test in recovered depressed patients. *J Abnorm Psychol, 110*(2), 282-289.

Gillath, Mikulincer, Fitzsimons, Shaver, Schachner, & Bargh. (2006). Automatic Activation of Attachment-Related Goals. *Personality and Social Psychology Bulletin, 32*(10), 1375-1388. doi: 10.1177/0146167206290339

Glashouwer, & de Jong. (2009). Disorder-specific automatic self-associations in depression and anxiety: results of The Netherlands Study of Depression and Anxiety. *Psychological Medicine, 40*(07), 1101-1111. doi: 10.1017/s0033291709991371

Glashouwer, de Jong, & Penninx. (2011). Predictive validity of automatic self-associations for the onset of anxiety disorders. *Journal of Abnormal Psychology, 120*(3), 607.

Glashouwer, Vroling, de Jong, Lange, & de Keijser. (2013). Low implicit self-esteem and dysfunctional automatic associations in social anxiety disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 44*(2), 262-270.

Gorenstein, Andrade, Tung, & Artes. (1999). Psychometric properties of the Portuguese version of the Beck Depression Inventory on Brazilian college students. *Journal of Clinical Psychology, 55*(5), 553-562.

Gotlib, & Joormann. (2010). Cognition and depression: current status and future directions. *Annual Review of Clinical Psychology, 6*, 285.

Gray, LaPlante, Bannon, Ambady, & Shaffer. (2011). Development and validation of the alcohol identity implicit associations test (AI-IAT). *Addictive Behaviors, 36*(9), 919-926.

Green, Furrer, & McAllister. (2007). How Do Relationships Support Parenting? Effects of Attachment Style and Social Support on Parenting Behavior in an At-Risk Population. *American Journal of Community Psychology, 40*(1-2), 96-108. doi: 10.1007/s10464-007-9127-y

Greenwald, & Banaji. (1995). Implicit Social Cognition: Attitudes, Self-Esteem, and Stereotypes. *Psychological Review, 102*(1), 4-27.

Greenwald, Banaji, Rudman, Farnham, Nosek, & Mellott. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*(1), 3-25. doi: 10.1037//0033-295x.109.1.3

Greenwald, & Farnham. (2000). Using the Implicit Association Test to measure Self-Esteem and Self-Concept. *Journal of Personality and Social Psychology, 79*(6), 1022-1038.

Greenwald, McGhee, & Schwartz. (1998). Measuring individual differences in implicit cognition: the implicit association test. *J Pers Soc Psychol, 74*(6), 1464-1480.

Greenwald, & Nosek. (2001). Health of the Implicit Association Test at age 3. *Zeitschrift für Experimentelle Psychologie, 48*(2), 85-93.

Greenwald, Nosek, & Banaji. (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, Poehlman, Uhlmann, & Banaji. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, Vol.97*(1), pp.

Gregg, & Sedikides. (2010). Narcissistic fragility: Rethinking its links to explicit and implicit self-esteem. *Self and Identity, 9*(2), 142-161.

Groenewold, Opmeer, de Jonge, Aleman, & Costafreda. (2013). Emotional valence modulates brain functional abnormalities in depression: evidence from a meta-analysis of fMRI studies. *Neuroscience & Biobehavioral Reviews, 37*(2), 152-163.

Gual, Pérez‑Gaspar, Martínez‑González, Lahortiga, Irala‑Estévez, & Cervera‑Enguix. (2002). Self‑esteem, personality, and eating disorders: Baseline assessment of a prospective population‑based cohort. *International Journal of Eating Disorders, 31*(3), 261-273.

Haeffel, Abramson, Brazy, Shah, Teachman, & Nosek. (2007). Explicit and implicit cognition: a preliminary test of a dual-process theory of cognitive vulnerability to depression. *Behav Res Ther, 45*(6), 1155-1167. doi: S0005-7967(06)00206-3 [pii]

10.1016/j.brat.2006.09.003

Hamilton. (1960). A rating scale for depression. *Journal of neurology, neurosurgery, and psychiatry, 23*(1), 56.

Hamilton, Etkin, Furman, Lemus, Johnson, & Gotlib. (2012). Functional neuroimaging of major depressive disorder: a meta-analysis and new integration of baseline activation and neural response data. *American Journal of Psychiatry, 169*(7), 693-703.

Hardeveld, Spijker, De Graaf, Nolen, & Beekman. (2010). Prevalence and predictors of recurrence of major depressive disorder in the adult population. *Acta Psychiatrica Scandinavica, 122*(3), 184-191.

Harmon-Jones, Simon, Greenberg, Pyszczynski, Solomon, & McGregor. (1997). Terror management theory and self-esteem: Evidence that increased self-esteem reduced mortality salience effects. *Journal of Personality and Social Psychology, 72*(1), 24.

Harrison. (2014). *The experience of stress in adolescents living in low-income communities in the Western Cape: The role of self-esteem, coping and perceived social support.* Stellenbosch: Stellenbosch University.

Harrison, Stritzke, Fay, Ellison, & Hudaib. (2014). Probing the implicit suicidal mind: Does the Death/Suicide Implicit Association Test reveal a desire to die, or a diminished desire to live? *Psychological Assessment, 26*(3), 831.

Hart, Whalen, Shin, McInerney, Fischer, & Rauch. (2000). Differential response in the human amygdala to racial outgroup vs ingroup face stimuli. *Neuroreport, 11*(11), 2351-2354.

Hayes. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*: Guilford Press.

Heatherton, & Polivy. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology, 60*(6), 895.

Heilbig, Putnam, Gartstein, & Rothbart. (2010). *Development and assessment of Short and Very Short forms of the Early Childhood Behavior Questionnaire*. Paper presented at the Biannual Convention of the International Society on Infant Studies, Baltimore.

Hertel. (2002). Cognitive biases in anxiety and depression: Introduction to the special issue. *Cognition & Emotion, 16*(3), 321-330.

Hetts, Sakuma, & Pelham. (1999). Two roads to positive regard: Implicit and explicit self-evaluation and culture. *Journal of Experimental Social Psychology, 35*(6), 512-559.

Hock, McBride, & Gnezda. (1989). Maternal Separation Anxiety: Mother-Infant Separation from the Maternal Perspective. *Child Development, 60*(4), 793-802. doi: 10.2307/1131019

Hofmann, Gawronski, Gschwendner, Le, & Schmitt. (2005a). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Pers Soc Psychol Bull, 31*(10), 1369-1385. doi: 31/10/1369 [pii]
10.1177/0146167205275613

Hofmann, Gschwendner, Nosek, & Schmitt. (2005b). What moderates implicit— explicit consistency? *European Review of Social Psychology, 16*(1), 335-390.

Holden, & Edwards. (1989). Parental Attitudes Toward Child Rearing: Instruments, Issues, and Implications. *Psychological Bulletin, 106*(1), 29-58.

Hoorens, & Nuttin. (1993). Overvaluation of own attributes: Mere ownership or subjective frequency? *Social Cognition, 11*(2), 177-200.

Howes, Galinsky, & Kontos. (1998). Child care caregiver sensitivity and attachment. *Social Development, 7*(1), 25-36.

Huijding, & de Jong. (2007). Beyond fear and disgust: the role of (automatic) contamination-related associations in spider phobia. *J Behav Ther Exp Psychiatry, 38*(2), 200-211. doi: 10.1016/j.jbtep.2006.10.009

Hunger, Sabariego, Stollenwerk, Cieza, & Leidl. (2012). Validity, reliability and responsiveness of the EQ-5D in German stroke patients undergoing rehabilitation. *Quality of Life Research, 21*(7), 1205-1216.

Hurst, Kind, Ruta, Hunter, & Stubbings. (1997). Measuring Health-Related Quality Of Life In Rheumatoid Arthritis: Validity, Responsiveness And Reliability Of Euroqol (Eq-5D). *British Journal of Rheumatology, 36*, 551-559.

Isabella. (1993). Origins of Attachment: Maternal Interactive Behavior across the First Year. *Child Development, 64*(2), 605-621. doi: 10.2307/1131272

Isabella, Belsky, & von Eye. (1989). Origins of infant-mother attachment: An examination of interactional synchrony during the infant's first year. *Developmental Psychology, 25*(1), 12.

Jabben, de Jong, Kupka, Glashouwer, Nolen, & Penninx. (2014). Implicit and explicit self-associations in bipolar disorder: A comparison with healthy controls and unipolar depressive disorder. *Psychiatry Research, 215*(2), 329-334.

James. (1984). *Psychology, briefer course* (Vol. 14): Harvard University Press.

Janis, & King. (1954). The influence of role playing on opinion change. *The Journal of Abnormal and Social Psychology, 49*(2), 211.

Janiszewski, & Wyer Jr. (2014). Content and process priming: A review. *Journal of Consumer Psychology, 24*(1), 96-118.

Johnstone, van Reekum, Urry, Kalin, & Davidson. (2007). Failure to regulate: counterproductive recruitment of top-down prefrontal-subcortical circuitry in major depression. *The Journal of Neuroscience, 27*(33), 8877-8884.

Jones, Pelham, Mirenberg, & Hetts. (2002). Name letter preferences are not merely mere exposure: Implicit egotism as self-regulation. *Journal of Experimental Social Psychology, 38*(2), 170-177.

Joormann. (2004). Attentional bias in dysphoria: The role of inhibitory processes. *Cognition and Emotion, 18*(1), 125-147.

Jordan, Logel, Spencer, Zanna, Wood, & Holmes. (2013). Responsive Low Self-Esteem: Low Explicit Self-Esteem, Implicit Self-Esteem, and Reactions to Performance Outcomes. *Journal of Social and Clinical Psychology, 32*(7), 703-732.

Jordan, Spencer, Zanna, Hoshino-Browne, & Correll. (2003). Secure and Defensive High Self-Esteem. *Journal of Personality and Social Psychology, 85*(5), 969-978. doi: 10.1037/0022-3514.85.5.969

Jordan, Whitfield, & Zeigler-Hill. (2007). Intuition and the correspondence between implicit and explicit self-esteem. *Journal of Personality and Social Psychology, 93*(6), 1067.

Kahneman. (2012, September 26). A proposal to deal with questions about priming effects [web log post]. from http://www.nature.com/polopoly_fs/7.6716.1349271308!/suppinfoFile/Kahneman Letter.pdf

Kandel. (1999). Biology and the future of psychoanalysis: a new intellectual framework for psychiatry revisited. *American Journal of Psychiatry, 156*(4), 505-524.

Karpinski, & Hilton. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology, 81*(5), 774.

Karpinski, & Steinberg. (2006). Implicit and Explicit Self-Esteem: Theoretical and Methodological Refinements. In M. H. Kernis (Ed.), *Self-esteem issues and answers: A sourcebook of current perspectives*. Hove: Psychology Press.

Katz, Allport, & Jenness. (1931). Students' attitudes; a report of the Syracuse University reaction study.

Kawakami, & Dovidio. (2001). The reliability of implicit stereotyping. *Personality and Social Psychology Bulletin, 27*(2), 212-225.

Kendler, Gardner, & Prescott. (2014a). Toward a comprehensive developmental model for major depression in men.

Kendler, Gardner, & Prescott. (2014b). Toward a comprehensive developmental model for major depression in women.

Keren, Dollberg, Koster, Danino, & Feldman. (2010). Family Functioning and Interactive Patterns in the Context of Infant Psychopathology. *Journal of Family Psychology, 24*(5), 597-604.

Keren, Feldman, & Tyano. (2001). Diagnoses and interactive patterns of infants referred to a community-based infant mental health clinic. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(1), 27-35.

Kernis. (2003). Toward a conceptualization of optimal self-esteem. *Psychological Inquiry, 14*(1), 1-26.

Kessler, Chiu, Demler, & Walters. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 617-627.

Kesting, Mehl, Rief, Lindenmeyer, & Lincoln. (2011). When paranoia fails to enhance self-esteem: Explicit and implicit self-esteem and its discrepancy in patients with persecutory delusions compared to depressed and healthy controls. *Psychiatry Research, 186*(2), 197-202.

Kiefer, & Sekaquaptewa. (2007). Implicit stereotypes, gender identification, and math-related outcomes a prospective study of female college students. *Psychological Science, 18*(1), 13-18.

Kihlstrom. (1987). The cognitive unconscious. *Science, 237*(4821), 1445-1452.

Kihlstrom. (1994). Commentary: Psychodynamics and social cognition—Notes on the fusion of psychoanalysis and psychology. *Journal of Personality, 62*(4), 681-696.

Kihlstrom, Barnhardt, & Tataryn. (1992). The Psychological Unconscious: Found, Lost, and Regained. *American Psychologist, 47*(6), 788-791.

Kihlstrom, Mulvaney, Tobias, & Tobis. (2000). The emotional unconscious. In E. Eich, J. F. Kihlstrom, G. Bower, J. Forgas & P. Niedental (Eds.), *Counterpoints: Cognition and emotion* (pp. 30-86). New York: Oxford University Press.

Kim. (2003). Voluntary controllability of the implicit association test (IAT). *Social Psychology Quarterly*, 83-96.

Kline. (2013). *Handbook of psychological testing*. London: Routledge.

Kling, Hyde, Showers, & Buswell. (1999). Gender differences in self-esteem: a meta-analysis. *Psychological Bulletin, 125*(4), 470.

Koestner, & Mageau. (2006). The Assessment of Implicit and Explicit Self-Esteem: Lessons from Motive Research. In M. H. Kernis (Ed.), *Self-esteem issues and answers: A sourcebook of current perspectives*. Hove: Psychology Press.

Kohut, & Wolf. (1978). The disorders of the self and their treatment: An outline. *International journal of Psychoanalysis, 59*(4), 413-425.

Koole. (2004). Volitional shielding of the self: Effects of action orientation and external demands on implicit self-evaluation. *Social Cognition, 22*(1: Special issue), 100-125.

Koole, DeHart, Nimrod, Fryer, Salvator, Knihnicki, . . . Habashi. (2007). Self-affection without self-reflection: Origins, models, and consequences of implicit self-esteem. *The self*, 21-49.

Kroenke, & Spitzer. (2002). The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric Annals, 32*(9), 1-7.

Kroenke, Spitzer, & Williams. (2001). The Phq‐9. *Journal of general internal medicine, 16*(9), 606-613.

Kruglanski, & Stroebe. (2012). *Handbook of the history of social psychology*: Psychology Press.

Kruijt, Antypa, Booij, de Jong, Glashouwer, Penninx, & Van der Does. (2013). Cognitive reactivity, implicit associations, and the incidence of depression: a two-year prospective study. *PLoS One, 8*(7), e70245.

Kuster, Orth, & Meier. (2012). Rumination mediates the prospective effect of low self-esteem on depression a five-wave longitudinal study. *Personality and Social Psychology Bulletin, 38*(6), 747-759.

Lane, Banaji, Nosek, & Greenwald. (2007). Understanding and using the implicit association test: IV. *Implicit measures of attitudes*, 59-102.

LaPiere. (1934). Attitudes vs. actions. *Social forces, 13*(2), 230-237.

Laplanche, & Pontalis. (1988). *The language of psychoanalysis*: Karnac Books.

Leary, Schreindorfer, & Haupt. (1995). The role of low self-esteem in emotional and behavioral problems: Why is low self-esteem dysfunctional? *Journal of Social and Clinical Psychology, 14*(3), 297-314.

Lemma, Target, & Fonagy. (2011). *Brief Dynamic Interpersonal Therapy. A clinician's guide*. Oxford: Oxford University Press.

Lemmens, Roefs, Arntz, van Teeseling, Peeters, & Huibers. (2014). The value of an implicit self-associative measure specific to core beliefs of depression. *Journal of Behavior Therapy and Experimental Psychiatry, 45*(1), 196-202.

Lewinsohn, Hoberman, & Rosenbaum. (1988). A prospective study of risk factors for unipolar depression. *Journal of Abnormal Psychology, 97*(3), 251.

Lima. (2007). The association between narcissism and implicit self-esteem: a test of the fragile self-esteem hypothesis.

Lohaus, Keller, Ball, Elben, & Voelker. (2001). Maternal Sensitivity: Components and Relations to Warmth and Contingency. *Parenting, 1*(4), 267-284. doi: 10.1207/s15327922par0104_1

Lorenzini, & Fonagy. (2013). Attachment and personality disorders: a short review. *FOCUS: The Journal of Lifelong Learning in Psychiatry, 11*(2), 155-166.

Luborsky, & Barrett. (2006). The History and Empirical Status of Key Psychoanalytic Concepts. *Annual Review of Clinical Psychology, 2*(1), 1-19. doi: doi:10.1146/annurev.clinpsy.2.022305.095328

Lutz, & Hock. (1995). Maternal Separation Anxiety: Relations to Adult Attachment Representations in Mothers of Infants. *The Journal of Genetic Psychology, 156*(1), 57-72. doi: 10.1080/00221325.1995.9914806

Luyten, Blatt, & Corveleyn. (2005). Epilogue: Towards integration in the theory and treatment of depression? The time is now. In J. Corveleyn, P. Luyten, S. J. Blatt & H. Lens-Gielis (Eds.), *The theory and treatment of depression: Towards a dynamic interactionism model* (pp. 253-284). Leuven: Leuven University Press.

Luyten, Mayes, Nijssens, & Fonagy. (2013). The Parental Reflective Functioning Questionnaire [Manuscript submitted for publication].

Maier, Bernier, Pekrun, Zimmermann, & Grossmann. (2004). Attachment working models as unconscious structures: An experimental test. *International Journal of Behavioral Development, 28*(2), 180-189. doi: 10.1080/01650250344000398

Main, Kaplan, & Cassidy. (1985). Security in infancy, childhood, and adulthood: A move to the level of representation. *Monographs of the Society fir Research in Child Development, 50*(1-2), 66-104. doi: 10.2307/3333827

Maison, Greenwald, & Bruin. (2001). The Implicit Association Test as a measure of implicit consumer attitudes.

Maison, Greenwald, & Bruin. (2004). Predictive validity of the Implicit Association Test in studies of brands, consumer attitudes, and behavior. *Journal of Consumer Psychology, 14*(4), 405-415.

Mandić-Gajić, Samardžić, & Špirić. (2015). Correlation and characteristics of self-rating and clinical rating of depression among alcoholics in the course of early abstinence. *Vojnosanitetski pregled*(00), 47-47.

Marhe, Waters, van de Wetering, & Franken. (2013). Implicit and explicit drug-related cognitions during detoxification treatment are associated with drug relapse: An ecological momentary assessment study. *Journal of Consulting and Clinical Psychology, 81*(1), 1.

Masling. (2000). Empirical evidence and the health of psychoanalysis. *Journal of the American Academy of Psychoanalysis, 28*, 665-685.

Masling, & Cohen. (1987). Psychotherapy, clinical evidence, and the self-fulfilling prophecy. *Psychoanalytic Psychology, 4*(1), 65-79.

Matt, Vázquez, & Campbell. (1992). Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clinical Psychology Review, 12*(2), 227-255.

Mayseless, & Scher. (2000). Mother's Attachment Concerns Regarding Spouse and Infant's Temperament as Modulators of Maternal Separation Anxiety. *Journal of Child Psychology and Psychiatry, 41*(7), 917-925. doi: 10.1111/1469-7610.00679

McCabe, & Gotlib. (1995). Selective attention and clinical depression: performance on a deployment-of-attention task. *Journal of Abnormal Psychology, 104*(1), 241.

McCabe, & Toman. (2000). Stimulus exposure duration in a deployment-of-attention task: Effects on dysphoric, recently dysphoric, and nondysphoric individuals. *Cognition & Emotion, 14*(1), 125-142.

McConnell, & Leibold. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37*(5), 435-442.

McDonald, Asher, Kerr, & Navarrete. (2011). Fertility and Intergroup Bias in Racial and Minimal-Group Contexts: Evidence for Shared Architecture. *Psychological Science, 22*(7), 860-865. doi: 10.1177/0956797611410985

McDowell. (2006). *Measuring health: a guide to rating scales and questionnaires*: Oxford University Press.

McGregor, Nail, Marigold, & Kang. (2005). Defensive pride and consensus: strength in imaginary numbers. *Journal of Personality and Social Psychology, 89*(6), 978.

Meins. (1997). *Security of attachment and the social development of congnition*. Hove: Psychology Press.

Meins, Fernyhough, Wainwright, Gupta, Fradley, & Tuckey. (2002). Maternal Mind-Mindedness and Attachment Security as Predictors of Theory of Mind Understanding. *Child Development, 73*(6), 1715-1726. doi: 10.2307/3696412

Meites, Deveney, Steele, Holmes, & Pizzagalli. (2008). Implicit depression and hopelessness in remitted depressed individuals. *Behaviour Research and Therapy, 46*(9), 1078-1084. doi: 10.1016/j.brat.2008.05.008

Melles, Kuile, Dewitte, Lankveld, Brauer, & Jong. (2014). Automatic and Deliberate Affective Associations with Sexual Stimuli in Women with Lifelong Vaginismus Before and After Therapist‐Aided Exposure Treatment. *The journal of sexual medicine, 11*(3), 786-799.

Metalsky, Joiner, Hardin, & Abramson. (1993). Depressive reactions to failure in a naturalistic setting: A test of the hopelessness and self-esteem theories of depression. *Journal of Abnormal Psychology, 102*(1), 101.

Mikulincer, Birnbaum, Woddis, & Nachmias. (2000). Stress and Accessibility of Proximity-Related Thoughts: Exploring the Normative and Intraindividual Components of Attachment Theory. *Journal of Personality & Social Psychology, 78*(3), 509-523.

Mikulincer, Gillath, & Shaver. (2002). Activation of the Attachment System in Adulthood: Threat-Related Primes Increase the Accessibility of Mental Representations of Attachment Figures. *Journal of Personality & Social Psychology, 83*(4), 881-895.

Mikulincer, Hirschberger, Nachmias, & Gillath. (2001). The Affective Component of the Secure Base Schema: Affective Priming With Representations of Attachment Security. *Journal of Personality & Social Psychology, 81*(2), 305-321.

Mikulincer, & Shaver. (2009). An attachment and behavioral systems perspective on social support. *Journal of Social and Personal Relationships, 26*(1), 7-19. doi: 10.1177/0265407509105518

Mikulincer, & Shaver. (2012). Adult Attachment Orientations and Relationship Processes. *Journal of Family Theory & Review, 4*(4), 259-274. doi: 10.1111/j.1756-2589.2012.00142.x

Mikulincer, Shaver, Sapir-Lavid, & Avihou-Kanza. (2009). What's Inside the Minds of Securely and Insecurely Attached People? The Secure-Base Script and Its Associations With Attachment-Style Dimensions. *Journal of Personality & Social Psychology, 97*(4), 615-633.

Millar. (2011). Predicting dental flossing behavior: The role of implicit and explicit responses and beliefs. *Basic and Applied Social Psychology, 33*(1), 7-15.

Modell. (2008). Implicit or Unconscious?: Commentary on Paper by the Boston Change Process Study Group. *Psychoanalytic Dialogues, 18*(2), 162-167. doi: 10.1080/10481880801909534

Monteith, Voils, & Ashburn-Nardo. (2001). Taking a look underground: Detecting, interpreting, and reacting to implicit racial biases. *Social Cognition, 19*(4), 395-417.

Moring, Bowen, & Thomas. (2014). Use of the Implicit Association Test for the Measurement of Tinnitus-Related Distress. *American journal of audiology, 23*(3), 293-302.

Moss, Dubois-Comtois, Cyr, Tarabulsy, St-Laurent, & Bernier. (2011). Efficacy of a home-visiting intervention aimed at improving maternal sensitivity, child attachment, and behavioral outcomes for maltreated children: A randomized control trial. *Development and Psychopathology, 23*(01), 195-210.

Mount, Crockenberg, Jó, & Wagar. (2010). Maternal and child correlates of anxiety in 2½-year-old children. *Infant Behavior and Development, 33*(4), 567-578.

Nakamura, Hayakawa, Okamura, Kohigashi, Fukui, & Narumoto. (2015). Defensive function of persecutory delusion and discrepancy between explicit and implicit self-esteem in schizophrenia: study using the Brief Implicit Association Test. *Neuropsychiatric Disease and Treatment, 11*, 33.

Nakonezny, Carmody, Morris, Kurian, & Trivedi. (2010). Psychometric evaluation of the Snaith–Hamilton Pleasure Scale (SHAPS) in adult outpatients with major depressive disorder. *International clinical psychopharmacology, 25*(6), 328-333. doi: 10.1097/YIC.0b013e32833eb5ee

Nier. (2005). How dissociated are implicit and explicit racial attitudes? A bogus pipeline approach. *Group Processes & Intergroup Relations, 8*(1), 39-52.

Nisbett, & Wilson. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review, 84*(3), 231.

Nock, & Banaji. (2007). Prediction of suicide ideation and attempts among adolescents using a brief performance-based test. *Journal of Consulting and Clinical Psychology, 75*(5), 707.

Nock, Park, Finn, Deliberto, Dour, & Banaji. (2010). Measuring the suicidal mind: implicit cognition predicts suicidal behavior. *Psychol Sci, 21*(4), 511-517. doi: 0956797610364762 [pii]

10.1177/0956797610364762

Nolen-Hoeksema. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology, 109*(3), 504.

Nosek, & Banaji. (2001). The go/no-go association task. *Social Cognition, 19*(6), 625-666.

Nosek, Banaji, & Greenwald. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice, 6*(1), 101.

Nosek, Greenwald, & Banaji. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Pers Soc Psychol Bull, 31*(2), 166-180. doi: 31/2/166 [pii]

10.1177/0146167204271418

Nosek, Greenwald, & Banaji. (2007a). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Automatic Processes in Social Thinking and Behaviour* (pp. 265-292): Psychology Press.

Nosek, & Smyth. (2007). A multitrait-multimethod validation of the Implicit Association Test: implicit and explicit attitudes are related but distinct constructs. *Exp Psychol, 54*(1), 14-29.

Nosek, Smyth, Hansen, Devos, Lindner, Ranganath, . . . Greenwald. (2007b). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology, 18*(1), 36-88.

Nuttin. (1985). Narcissism beyond Gestalt and awareness: The name letter effect. *European Journal of Social Psychology, 15*(3), 353-361.

Oates, & Gervai. (2000). *Mothers' models of their infants [Unpublished work]*. Child and Youth Studies Group, The Open University and the Institute for Psychology of the Hungarian Academy of Sciences. Milton Keynes and Budapest.

Olson, & Fazio. (2001). Implicit attitude formation through classical conditioning. *Psychological Science, 12*(5), 413-417.

Olson, & Fazio. (2002). Implicit acquisition and manifestation of classically conditioned attitudes. *Social Cognition, 20*(2), 89-104.

Olson, & Fazio. (2003). Relations Between Implicit Measures of Prejudice. What Are We Measuring? *Psychological Science, 14*(6), 636-639.

Olson, Fazio, & Hermann. (2007). Reporting tendencies underlie discrepancies between implicit and explicit measures of self-esteem. *Psychological Science, 18*(4), 287-291.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716.

Orth, & Robins. (2013). Understanding the link between low self-esteem and depression. *Current Directions in Psychological Science, 22*(6), 455-460.

Orth, Robins, Meier, & Conger. (2015). Refining the Vulnerability Model of Low Self-Esteem and Depression: Disentangling the Effects of Genuine Self-Esteem and Narcissism.

Orth, Robins, & Widaman. (2012). Life-span development of self-esteem and its effects on important life outcomes. *Journal of Personality and Social Psychology, 102*(6), 1271.

Osman, Downs, Barrios, Kopper, Gutierrez, & Chiros. (1997). Factor structure and psychometric characteristics of the Beck Depression Inventory-II. *Journal of Psychopathology and Behavioral Assessment, 19*(4), 359-376.

Park, Ryu, Ha, Lee, Choi, Ha, & Cho. (2014). Assessment of implicit self-esteem in bipolar manic and euthymic patients using the implicit association test. *Comprehensive Psychiatry, 55*(3), 557-564.

Patten, Li Wang, Williams, Currie, Beck, Maxwell, & el-Guebaly. (2006). Descriptive epidemiology of major depression in Canada. *Canadian Journal of Psychiatry, 51*(2), 84.

Pavlickova, Turnbull, & Bentall. (2014). Discrepancies between explicit and implicit self‑esteem and their relationship to symptoms of depression and mania. *Psychology and Psychotherapy: Theory, Research and Practice, 87*(3), 311-323.

Pederson, & Moran. (1995). A Categorical Description of Infant-Mother Relationships in the Home and Its Relation to Q-Sort Measures of Infant-Mother Interaction. *Monographs of the Society for Research in Child Development, 60*(2/3), 111-132. doi: 10.2307/1166174

Pelham, Koole, Hardin, Hetts, Seah, & DeHart. (2005). Gender moderates the relation between implicit and explicit self-esteem. *Journal of Experimental Social Psychology, 41*(1), 84-89.

Pelham, Mirenberg, & Jones. (2002). Why Susie sells seashells by the seashore: implicit egotism and major life decisions. *Journal of Personality and Social Psychology, 82*(4), 469.

Phelps, & LeDoux. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron, 48*(2), 175-187.

Phelps, O'Connor, Cunningham, Funayama, Gatenby, Gore, & Banaji. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of cognitive neuroscience, 12*(5), 729-738.

Phillips, & Hine. (2014). En Route to Depression: Self‑Esteem Discrepancies and Habitual Rumination. *Journal of Personality*.

Phillips, Hine, & Bhullar. (2012). A Latent Profile Analysis of Implicit and Explicit Cognitions Associated with Depression. *Cognitive Therapy and Research, 36*(5), 458-473.

Phillips, Hine, & Thorsteinsson. (2010). Implicit cognition and depression: A meta-analysis. *Clinical Psychology Review, 30*(6), 691-709. doi: http://dx.doi.org/10.1016/j.cpr.2010.05.002

Picardi, Toni, & Caroppo. (2005). Stability of alexithymia and its relationships with the 'big five' factors, temperament, character, and attachment style. *Psychotherapy and Psychosomatics, 74*(6), 371-378.

Price, Nock, Charney, & Mathew. (2009). Effects of Intravenous Ketamine on Explicit and Implicit Measures of Suicidality in Treatment-Resistant Depression. *Biological Psychiatry, 66*(5), 522-526. doi: 10.1016/j.biopsych.2009.04.029

Putnam, Helbig, Gartstein, Rothbart, & Leerkes. (2014). Development and assessment of short and very short forms of the Infant Behavior Questionnaire–Revised. *Journal of Personality Assessment, 96*(4), 445-458.

Pyszczynski, & Greenberg. (1987). Self-regulatory perseveration and the depressive self-focusing style: a self-awareness theory of reactive depression. *Psychological Bulletin, 102*(1), 122.

Pyszczynski, Greenberg, Solomon, Arndt, & Schimel. (2004). Why do people need self-esteem? A theoretical and empirical review. *Psychological Bulletin, 130*(3), 435.

Randall, Rowe, Dong, Nock, & Colman. (2013). Assessment of self-harm risk using implicit thoughts. *Psychological Assessment, 25*(3), 714.

Remue, De Houwer, Barnes-Holmes, Vanderhasselt, & De Raedt. (2013). Self-esteem revisited: Performance on the implicit relational assessment procedure as a

measure of self-versus ideal self-related cognitions in dysphoria. *Cognition & Emotion, 27*(8), 1441-1449.

Remue, Hughes, De Houwer, & De Raedt. (2014). To Be or Want to Be: Disentangling the Role of Actual versus Ideal Self in Implicit Self-Esteem. *PLoS One, 9*(9), e108837. doi: 10.1371/journal.pone.0108837

Ritter, Ertel, Beil, Steffens, & Stangier. (2013). In the presence of social threat: Implicit and explicit self-esteem in social anxiety disorder. *Cognitive Therapy and Research, 37*(6), 1101-1109.

Robins, & Trzesniewski. (2005). Self-Esteem Development Across the Lifespan. *Current Directions in Psychological Science, 14*(3), 158-162. doi: 10.1111/j.0963-7214.2005.00353.x

Robins, Trzesniewski, Tracy, Gosling, & Potter. (2002). Global self-esteem across the life span. *Psychology and aging, 17*(3), 423.

Roefs, Huijding, Smulders, MacLeod, de Jong, Wiers, & Jansen. (2011). Implicit measures of association in psychopathology research. *Psychological Bulletin, 137*(1), 149.

Roefs, & Jansen. (2002). Implicit and explicit attitudes toward high-fat foods in obesity. *Journal of Abnormal Psychology, 111*(3), 517.

Rosenberg. (1965). *Society and the adolescent self-image* (P. U. Press Ed.). Princeton, NJ.

Roth, Steffens, Morina, & Stangier. (2012). Changed for the worse: subjective change in implicit and explicit self-esteem in individuals with current, past, and no posttraumatic stress disorder. *Psychotherapy and Psychosomatics, 81*(1), 64.

Rothermund, & Brandstädter. (2003). Coping with deficits and losses in later life: from compensatory action to accommodation. *Psychology and aging, 18*(4), 896.

Rude, Wenzlaff, Gibbs, Vane, & Whitney. (2002). Negative processing biases predict subsequent depressive symptoms. *Cognition & Emotion, 16*(3), 423-440.

Rudman. (2008). The validity of the Implicit Association Test is a scientific certainty. *Industrial and Organizational Psychology: Perspectives on Science and Practice, Vol.1*(4), pp.

Rudman, Dohn, & Fairchild. (2007). Implicit self-esteem compensation: Automatic threat defense. *Journal of Personality and Social Psychology, 93*(5), 798.

Rudman, & Goodwin. (2004). Gender differences in automatic in-group bias: why do women like women more than men like men? *Journal of Personality and Social Psychology, 87*(4), 494.

Rudman, Greenwald, Mellott, & Schwartz. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition, 17*(4), 437-465.

Rudolph, Schröder-Abe, Schutz, Gregg, & Sedikides. (2008). Through a Glass, Less Darkly?: Reassessing Convergent and Discriminant Validity in Measures of Implicit Self-Esteem. *European Journal of Psychological Assessment, 24*(4), 273-281.

Rüsch, Corrigan, Todd, & Bodenhausen. (2010). Implicit self-stigma in people with mental illness. *The Journal of Nervous and Mental Disease, 198*(2), 150-153.

Rutland, Cameron, Milne, & McGeorge. (2005). Social Norms and Self‐Presentation: Children's Implicit and Explicit Intergroup Attitudes. *Child Development, 76*(2), 451-466.

Sacher, Neumann, Fünfstück, Soliman, Villringer, & Schroeter. (2012). Mapping the depressed brain: a meta-analysis of structural and functional alterations in major depressive disorder. *Journal of Affective Disorders, 140*(2), 142-148.

Salman Akhtar, & Thomson Jr. (1982). Overview: Narcissistic personality disorder. *Am J Psychiatry, 139*(1).

Sandstrom, & Jordan. (2008). Defensive self-esteem and aggression in childhood. *Journal of Research in Personality, 42*(2), 506-514.

Schaefer, & Edgerton. (1985). Parent and child correlates of parental modernity. In I. E. Sigel (Ed.), *Parental belief systems: the psychological consequences for children* (pp. 287-318). Hillsdale: Erlbaum.

Scher, Ingram, & Segal. (2005). Cognitive reactivity and vulnerability: Empirical evaluation of construct activation and cognitive diatheses in unipolar depression. *Clinical Psychology Review, 25*(4), 487-510.

Schmukle, & Egloff. (2004). Does the Implicit Association Test for assessing anxiety measure trait and state variance? *European Journal of Personality, 18*(6), 483-494.

Schnabel, Asendorpf, & Greenwald. (2008). Using Implicit Association Tests for the assessment of implicit personality self-concept. *Handbook of personality theory and testing, 2*.

Schnabel, Banse, & Asendorpf. (2006). Employing automatic approach and avoidance tendencies for the assessment of implicit personality self-concept: The Implicit Association Procedure (IAP). *Experimental Psychology, 53*(1), 69-76.

Schoenmaker, Juffer, van IJzendoorn, Linting, van der Voort, & Bakermans-Kranenburg. (2015). From maternal sensitivity in infancy to adult attachment representations: a longitudinal adoption study with secure base scripts. *Attachment & Human Development*(ahead-of-print), 1-16.

Schreiber, Bohn, Aderka, Stangier, & Steil. (2012). Discrepancies between implicit and explicit self-esteem among adolescents with social anxiety disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 43*(4), 1074-1081.

Schröder‐Abé, Rudolph, & Schütz. (2007a). High implicit self‐esteem is not necessarily advantageous: Discrepancies between explicit and implicit self‐esteem and their relationship with anger expression and psychological health. *European Journal of Personality, 21*(3), 319-339.

Schröder‐Abé, Rudolph, Wiesner, & Schütz. (2007b). Self‐esteem discrepancies and defensive reactions to social feedback. *International Journal of Psychology, 42*(3), 174-183.

Seymour, Giallo, Cooklin, & Dunning. (2014). Maternal anxiety, risk factors and parenting in the first post‐natal year. *Child: care, health and development*.

Shah, Fonagy, & Stratharn. (2010). Is attachment transmitted across generations? The plot thickens. *Clinical Child Psychology and Psychiatry, 15*(3), 329-345. doi: 10.1177/1359104510365449

Shanks, Newell, Lee, Balakrishnan, & Ekelund. (2013). Priming Intelligent Behavior: An Elusive Phenomenon. *PLoS One, 8*(4), e56515.

Shaver, & Mikulincer. (2002a). Attachment-related psychodynamics. *Attachment & Human Development, 4*(2), 133-161. doi: 10.1080/14616730210154171

Shaver, & Mikulincer. (2002b). Dialogue on adult attachment: Diversity and integration. *Attachment & Human Development, 4*(2), 243-257. doi: 10.1080/14616730210157484

Sherif. (1935). An experimental study of stereotypes. *The Journal of Abnormal and Social Psychology, 29*(4), 371.

Shin, Park, Ryu, & Seomun. (2008). Maternal sensitivity: A concept analysis. *Journal of Advanced Nursing, 64*(3), 304-314.

Siegle, Thompson, Carter, Steinhauer, & Thase. (2007). Increased amygdala and decreased dorsolateral prefrontal BOLD responses in unipolar depression: related and independent features. *Biological Psychiatry, 61*(2), 198-209.

Simon. (1990). The effects of an irrelevant directional cue on human information processing.

Simon, & Rudell. (1967). Auditory SR compatibility: the effect of an irrelevant cue on information processing. *Journal of applied psychology, 51*(3), 300.

Smith, & Nosek. (2011). Affective focus increases the concordance between implicit and explicit attitudes. *Social Psychology, 42*(4), 300.

Snaith, Hamilton, Morley, Humayan, Hargreaves, & Trigwell. (1995). A scale for the assessment of hedonic tone the Snaith-Hamilton Pleasure Scale. *The British Journal of Psychiatry, 167*(1), 99-103.

Sonntag, Konnopka, Leichsenring, Salzer, Beutel, Herpertz, . . . Nolting. (2013). Reliability, validity and responsiveness of the EQ-5D in assessing and valuing health status in patients with social phobia. *HEALTH QUAL LIFE OUT, 11*, 215.

Sowislo, & Orth. (2013). Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. *Psychological Bulletin, 139*(1), 213.

Spalding, & Hardin. (2000). Unconscious unease and self-handicapping: behavioural consequences of individual differences in implicit and explicit self-esteem. *Psychological Science, 10*, 535-539.

Sroufe. (1983). *Infant-caregiver attachment and patterns of adaptation in preschool: The roots of maladaptation and competence.* Paper presented at the Minnesota symposium in child psychology.

Stanley, Phelps, & Banaji. (2008). The Neural Basis of Implicit Attitudes. *Current Directions in Psychological Science, 17*(2), 164-170. doi: 10.1111/j.1467-8721.2008.00568.x

Stark, Reitmeir, Leidl, & König. (2010). Validity, reliability, and responsiveness of the EQ‐5D in inflammatory bowel disease in Germany. *Inflammatory Bowel Diseases, 16*(1), 42-51.

Steer, Ball, & Ranieri. (1999). Dimensions of the Beck Depression Inventory-II in clinically depressed outpatients. *Journal of Clinical Psychology, 55*(1), 117-128.

Steffens. (2004). Is the implicit association test immune to faking? *Exp Psychol, 51*(3), 165-179.

Steinberg. (2006). *Implicit and explicit self-esteem level and reactivity as predictors of depression.* (Doctoral thesis). Retrieved from http://search.proquest.com/docview/304980161/abstract/141BC9B90A82FBFD B3A/1?accountid=14511 ProQuest Dissertation and Theses database.

Steinberg, Karpinski, & Alloy. (2007). The exploration of implicit aspects of self-esteem in vulnerability-stress models of depression. *Self and Identity, 6*(2-3), 101-117. doi: 10.1080/15298860601118884

Sterling. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa. *Journal of the American statistical association, 54*(285), 30-34.

Sterling, Rosenbaum, & Weinkam. (1995). Publication decisions revisited: The effect of the outcome of statistical tests on the decision to publish and vice versa. *The American Statistician, 49*(1), 108-112.

Stevenson-Hinde, Chicot, Shouldice, & Hinde. (2013). Maternal anxiety, maternal sensitivity, and attachment. *Attachment & Human Development, 15*(5-6), 618-636.

Stevenson-Hinde, Shouldice, & Chicot. (2011). Maternal anxiety, behavioral inhibition, and attachment. *Attachment & Human Development, 13*(3), 199-215.

Stieger, Formann, & Burger. (2011). Humor styles and their relationship to explicit and implicit self-esteem. *Personality and Individual Differences, 50*(5), 747-750.

Strack, & Deutsch. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*(3), 220-247.

Suzuki, Sakurai, Yasuda, Harai, Kitamura, Takahashi, & Furukawa. (2003). Reliability, validity and standardization of the Japanese version of the Social Adjustment Scale‐Self Report. *Psychiatry and clinical neurosciences, 57*(4), 441-446.

Tafarodi. (1998). Paradoxical self-esteem and selectivity in the processing of social information. *Journal of Personality and Social Psychology, 74*(5), 1181.

Tafarodi. (2006). Moral value, agency, and the measurement of self-esteem. In M. H. Kernis (Ed.), *Self-esteem issues and answers: A sourcebook of current perspectives*. Hove: Psychology Press.

Tafarodi, & Milne. (2002). Decomposing global self‐esteem. *Journal of Personality, 70*(4), 443-484.

Tajfel. (1974). Social identity and intergroup behaviour. *Social Science Information/sur les sciences sociales*.

Tanner, Stopa, & De Houwer. (2006). Implicit views of the self in social anxiety. *Behaviour Research and Therapy, 44*(10), 1397-1409.

Tarabulsy, Bernier, Provost, Maranda, Larose, Moss, . . . Tessier. (2005). Another look inside the gap: ecological contributions to the transmission of attachment in a sample of adolescent mother-infant dyads. *Developmental Psychology, 41*(1), 212.

Teachman. (2005). Information processing and anxiety sensitivity: Cognitive vulnerability to panic reflected in interpretation and memory biases. *Cognitive Therapy and Research, 29*(4), 479-499.

Teachman, & Clerkin. (2007). Obsessional beliefs and the implicit and explicit morality of intrusive thoughts. *Cognition and Emotion, 21*(5), 999-1024.

Teachman, Gregg, & Woody. (2001). Implicit associations for fear-relevant stimuli among individuals with snake and spider fears. *Journal of Abnormal Psychology, 110*(2), 226.

Teachman, Marker, & Smith-Janik. (2008). Automatic associations and panic disorder: Trajectories of change over the course of treatment. *Journal of Consulting and Clinical Psychology, 76*(6), 988.

Teachman, Smith-Janik, & Saporito. (2007). Information processing biases and panic disorder: Relationships among cognitive and symptom measures. *Behaviour Research and Therapy, 45*(8), 1791-1811.

Teachman, & Woody. (2003). Automatic processing in spider phobia: Implicit fear associations over the course of treatment. *Journal of Abnormal Psychology, 112*(1), 100-109. doi: 10.1037/0021-843x.112.1.100

Teachman, Woody, & Magee. (2006). Implicit and explicit appraisals of the importance of intrusive thoughts. *Behaviour Research and Therapy, 44*(6), 785-805.

Teasdale, Moore, Hayhurst, Pope, Williams, & Segal. (2002). Metacognitive awareness and prevention of relapse in depression: empirical evidence. *Journal of Consulting and Clinical Psychology, 70*(2), 275.

Teige, Schnabel, Banse, & Asendorpf. (2004). Assessment of multiple implicit self‐concept dimensions using the Extrinsic Affective Simon Task (EAST). *European Journal of Personality, 18*(6), 495-520.

Tennen, & Herzberger. (1987). Depression, self-esteem, and the absence of self-protective attributional biases. *Journal of Personality and Social Psychology, 52*(1), 72.

Thorndike. (1920). A constant error in psychological ratings. *Journal of applied psychology, 4*(1), 25-29.

Thurstone. (1931). The measurement of social attitudes. *The Journal of Abnormal and Social Psychology, 26*(3), 249.

Thurstone, & Chave. (1929). The measurement of attitude.

Thurstone, & Peterson. (1933). Motion pictures and the social attitudes of children. *Motion pictures and youth. New York: Macmillan.*

Thush, & Wiers. (2007). Explicit and implicit alcohol-related cognitions and the prediction of future drinking in adolescents. *Addictive Behaviors, 32*(7), 1367-1383. doi: http://dx.doi.org/10.1016/j.addbeh.2006.09.011

Thush, Wiers, Ames, Grenard, Sussman, & Stacy. (2008). Interactions between implicit and explicit cognition and working memory capacity in the prediction of alcohol use in at-risk adolescents. *Drug and Alcohol Dependence, 94*(1–3), 116-124. doi: http://dx.doi.org/10.1016/j.drugalcdep.2007.10.019

Tulving, & Schacter. (1990). Priming and human memory systems. *Science, 247*(4940), 301-306.

Tulving, Schacter, & Stark. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*(4), 336.

Valiente, Cantero, Vázquez, Sanchez, Provencio, & Espinosa. (2011). Implicit and explicit self-esteem discrepancies in paranoia and depression. *Journal of Abnormal Psychology, 120*(3), 691.

Van Ijzendoorn. (1995). Adult Attachment Representations, Parental Responsiveness, and Infant Attachment: A Meta-Analysis on the Predictive Validity of the Adult Attachment Interview. *Psychological Bulletin, 117*(3), 387-403.

Van Ijzendoorn, Moran, Belsky, Pederson, Bakermans-Kranenburg, & Kneppers. (2000). The similarity of sibling's attachments to their mother. *Child Development, 71*(4), 1086-1098.

van Tuijl, de Jong, Sportel, de Hullu, & Nauta. (2014). Implicit and explicit self-esteem and their reciprocal relationship with symptoms of depression and social anxiety: a longitudinal study in adolescents. *Journal of Behavior Therapy and Experimental Psychiatry, 45*(1), 113-121.

Vartanian, Polivy, & Herman. (2004). Implicit cognitions and eating disorders: Their application in research and treatment. *Cognitive and Behavioral Practice, 11*(2), 160-167.

Vater, Ritter, Schröder-Abé, Schütz, Lammers, Bosson, & Roepke. (2013). When grandiosity and vulnerability collide: Implicit and explicit self-esteem in patients with narcissistic personality disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 44*(1), 37-47. doi: http://dx.doi.org/10.1016/j.jbtep.2012.07.001

Vater, Schröder-Abé, Schütz, Lammers, & Roepke. (2010). Discrepancies between explicit and implicit self-esteem are linked to symptom severity in borderline personality disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 41*(4), 357-364.

Veletanlic. (2007). *Assessing attachment models using the implicit association test.* (Bachelor in Psychology), Kristianstad University College, Kristianstad.

Verkuyten. (1996). Personal self-esteem and prejudice among ethnic majority and minority youth. *Journal of Research in Personality, 30*(2), 248-263.

Verkuyten, & Masson. (1995). 'New racism', self-esteem, and ethnic relations among minority and majority youth in the Netherlands. *Social Behavior and Personality: an international journal, 23*(2), 137-154.

Victoria, & Fontenelle. (2011). Teste de Associação Implícita no transtorno obsessivo-compulsivo (TAI-TOC): consistência interna e correlações com medidas explícitas. *Rev Psiq Clín, 38*(2), 53-56.

Waller, & Scheidt. (2006). Somatoform disorders as disorders of affect regulation: a development perspective. *International Review of Psychiatry, 18*(1), 13-24.

Wanke, Plessner, Gartner, & Friese. (2002). Measuring implicit consumer attitudes and predicting brand choice. *Advances in Consumer research, 29*(1), 222.

Watson, Little, Sawrie, & Biderman. (1992). Measures of the narcissistic personality: Complexity of relationships with self-esteem and empathy. *Journal of Personality Disorders, 6*(4), 434-449.

Wegener, Geiser, Alfter, Mierke, Imbierowicz, Kleiman, . . . Conrad. (2015). Changes of explicitly and implicitly measured self-esteem in the treatment of major depression: Evidence for implicit self-esteem compensation. *Comprehensive Psychiatry, 58*, 57-67.

Weissman, Merikangas, Priya, Kidd, Prusoff, Leckman, & Pauls. (1986). Understanding the clinical heterogeneity of major depression using family data. *Archives of General Psychiatry, 43*(5), 430-434.

Weissman, Prusoff, Thompson, Harding, & Myers. (1978). Social adjustment by self-report in a community sample and in psychiatric outpatients. *The Journal of Nervous and Mental Disease, 166*(5), 317-326.

Wenzlaff, & Bates. (1998). Unmasking a cognitive vulnerability to depression: how lapses in mental control reveal depressive thinking. *Journal of Personality and Social Psychology, 75*(6), 1559.

Wenzlaff, & Eisenberg. (2002). Mental control after dysphoria: Evidence of a suppressed, depressive bias. *Behavior therapy, 32*(1), 27-45.

Westen. (1998). The scientific legacy of Sigmund Freud: Toward a psychodynamically informed psychological science. *Psychological Bulletin, 124*(3), 333.

Whitley. (1978). The organization of scientific work in "configurational" and "restricted" sciences: a study of three research laboratories. *International Journal of Sociology, 8*(1/2), 95-112. doi: 10.2307/40783894

Wicker. (1969). Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social issues, 25*(4), 41-78.

Williams, Power, Millar, Freeman, Yellowlees, Dowds, . . . Jackson. (1993). Comparison of eating disorders and other dietary/weight groups on measures of perceived control, assertiveness, self‐esteem, and self‐directed hostility. *International Journal of Eating Disorders, 14*(1), 27-32.

Wills. (1981). Downward comparison principles in social psychology. *Psychological Bulletin, 90*(2), 245.

Wilson. (1968). Perceptual distortion of height as a function of ascribed academic status. *The Journal of Social Psychology, 74*(1), 97-102.

Wilson, Lindsey, & Schooler. (2000). A model of dual attitudes. *Psychological Review, 107*(1), 101.

Wisco. (2009). Depressive cognition: Self-reference and depth of processing. *Clinical Psychology Review, 29*(4), 382-392.

Wittchen, Jacobi, Rehm, Gustavsson, Svensson, Jönsson, . . . Faravelli. (2011). The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European Neuropsychopharmacology, 21*(9), 655-679.

Wolff, & Ijzendoorn. (1997). Sensitivity and attachment: A meta‐analysis on parental antecedents of infant attachment. *Child Development, 68*(4), 571-591.

Wouters, Duriez, Luyckx, Klimstra, Colpin, Soenens, & Verschueren. (2013). Depressive symptoms in university freshmen: Longitudinal relations with contingent self-esteem and level of self-esteem. *Journal of Research in Personality, 47*(4), 356-363.

Yamaguchi, Greenwald, Banaji, Murakami, Chen, Shiomura, . . . Krendl. (2007). Apparent universality of positive implicit self-esteem. *Psychological Science, 18*(6), 498-500.

Zajonc. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*(2p2), 1.

Zayas, & Shoda. (2004). Using the Implicit Association Test (IAT) to assess automatic reactions elicited by one's mother [Unpublished technical report]. Washington: University of Washington.

Zayas, & Shoda. (2005). Do Automatic Reactions Elicited by Thoughts of Romantic Partner, Mother, and Self Relate to Adult Romantic Attachment? *Personality and Social Psychology Bulletin, 31*(8), 1011-1025. doi: 10.1177/0146167204274100

Zeigler-Hill. (2006). Discrepancies Between Implicit and Explicit Self-Esteem: Implications for Narcissism and Self-Esteem Instability. *Journal of Personality, 74*(1), 119-144. doi: 10.1111/j.1467-6494.2005.00371.x

Zeigler-Hill, & Terry. (2007). Perfectionism and explicit self-esteem: The moderating role of implicit self-esteem. *Self and Identity, 6*(2-3), 137-153.

Zimmerman, Copeland, Shope, & Dielman. (1997). A Longitudinal Study of Self-Esteem: Implications for Adolescent Development. *Journal of Youth and Adolescence, 26*(2), 117-141. doi: 10.1023/A:1024596313925

Zweig, & Turkel. (2007). The social adjustment scale-self-report: psychometric properties for older adults 1, 2. *Psychological Reports, 101*(3), 920-926.

*Improved D calculation adapted from Greenwald, A. G., et al. (2003). "Understanding and using the Implicit Association Test: I. An improved scoring algorithm." *Journal of Personality and Social Psychology 85*(2): 197-216.

```
USE ALL.

execute.


DO REPEAT

Time =t1 t2 t3 t4 t5 t6 t7 t8 t9 t10 t11 t12 t13 t14 t15 t16 t17
t18 t19 t20 t21 t22 t23 t24 t25 t26 t27 t28 t29 t30 t31 t32 t33
t34 t35 t36 t37 t38 t39 t40 t41 t42 t43 t44 t45 t46 t47 t48 t49
t50 t51 t52 t53 t54 t55 t56 t57 t58 t59 t60 t61 t62 t63 t64 t65
t66 t67 t68 t69 t70 t71 t72 t73 t74 t75 t76 t77 t78 t79 t80 t81
t82 t83 t84 t85 t86 t87 t88 t89 t90 t91 t92 t93 t94 t95 t96 t97
t98 t99 t100 t101 t102 t103 t104 t105 t106 t107 t108 t109 t110
t111 t112 t113 t114 t115 t116 t117 t118 t119 t120/Correct = c1
c2 c3 c4 c5 c6 c7 c8 c9 c10 c11 c12 c13 c14 c15 c16 c17 c18 c19
c20 c21 c22 c23 c24 c25 c26 c27 c28 c29 c30 c31 c32 c33 c34 c35
c36 c37 c38 c39 c40 c41 c42 c43 c44 c45 c46 c47 c48 c49 c50 c51
c52 c53 c54 c55 c56 c57 c58 c59 c60 c61 c62 c63 c64 c65 c66 c67
c68 c69 c70 c71 c72 c73 c74 c75 c76 c77 c78 c79 c80 c81 c82 c83
c84 c85 c86 c87 c88 c89 c90 c91 c92 c93 c94 c95 c96 c97 c98 c99
c100 c101 c102 c103 c104 c105 c106 c107 c108 c109 c110 c111 c112
c113 c114 c115 c116 c117 c118 c119 c120/

TC = tcorrected1 to tcorrected120.

if (Correct = 1) TC = Time.

if (Correct=0) tc= 0.

end repeat.

EXECUTE.
```

*Creating a variable containing only correct responses and responses below 10,000 ms. to subsequently compute means for correct responses.

* While tcorrected variabes are those correct, upTC variables are correct minus those longer than 10,000.

RECODE tcorrected1 tcorrected2 tcorrected3 tcorrected4
tcorrected5 tcorrected6 tcorrected7 tcorrected8 tcorrected9
tcorrected10 tcorrected11 tcorrected12 tcorrected13 tcorrected14
tcorrected15 tcorrected16 tcorrected17 tcorrected18 tcorrected19
tcorrected20 tcorrected21 tcorrected22 tcorrected23 tcorrected24
tcorrected25 tcorrected26 tcorrected27 tcorrected28 tcorrected29
tcorrected30

tcorrected31 tcorrected32 tcorrected33 tcorrected34 tcorrected35
tcorrected36 tcorrected37 tcorrected38 tcorrected39 tcorrected40
tcorrected41 tcorrected42 tcorrected43 tcorrected44 tcorrected45
tcorrected46 tcorrected47 tcorrected48 tcorrected49 tcorrected50
tcorrected51 tcorrected52 tcorrected53 tcorrected54 tcorrected55
tcorrected56 tcorrected57 tcorrected58 tcorrected59 tcorrected60
tcorrected61 tcorrected62 tcorrected63 tcorrected64 tcorrected65
tcorrected66 tcorrected67 tcorrected68 tcorrected69 tcorrected70
tcorrected71 tcorrected72 tcorrected73 tcorrected74 tcorrected75
tcorrected76 tcorrected77 tcorrected78 tcorrected79 tcorrected80
tcorrected81 tcorrected82 tcorrected83 tcorrected84 tcorrected85
tcorrected86 tcorrected87 tcorrected88 tcorrected89 tcorrected90
tcorrected91 tcorrected92 tcorrected93 tcorrected94 tcorrected95
tcorrected96 tcorrected97 tcorrected98 tcorrected99
tcorrected100 tcorrected101 tcorrected102 tcorrected103
tcorrected104 tcorrected105 tcorrected106 tcorrected107
tcorrected108 tcorrected109 tcorrected110 tcorrected111
tcorrected112 tcorrected113 tcorrected114 tcorrected115
tcorrected116 tcorrected117 tcorrected118 tcorrected119
tcorrected120 (1 thru 10000=Copy) (ELSE=SYSMIS) INTO upTC1 upTC2
upTC3 upTC4 upTC5 upTC6 upTC7 upTC8 upTC9 upTC10 upTC11 upTC12

upTC13 upTC14 upTC15 upTC16 upTC17 upTC18 upTC19 upTC20 upTC21
upTC22 upTC23 upTC24 upTC25 upTC26 upTC27 upTC28 upTC29 upTC30
upTC31 upTC32 upTC33 upTC34 upTC35 upTC36 upTC37 upTC38 upTC39
upTC40 upTC41 upTC42 upTC43 upTC44 upTC45 upTC46 upTC47 upTC48
upTC49 upTC50 upTC51 upTC52 upTC53 upTC54 upTC55 upTC56 upTC57
upTC58 upTC59 upTC60 upTC61 upTC62 upTC63 upTC64 upTC65 upTC66
upTC67 upTC68 upTC69 upTC70 upTC71 upTC72 upTC73 upTC74 upTC75
upTC76 upTC77 upTC78 upTC79 upTC80 upTC81 upTC82 upTC83 upTC84
upTC85 upTC86 upTC87 upTC88 upTC89 upTC90 upTC91 upTC92 upTC93
upTC94 upTC95 upTC96 upTC97 upTC98 upTC99 upTC100 upTC101
upTC102 upTC103 upTC104 upTC105 upTC106 upTC107 upTC108 upTC109
upTC110 upTC111 upTC112 upTC113 upTC114 upTC115 upTC116 upTC117
upTC118 upTC119 upTC120.

EXECUTE.

*Compute mean of correct latencies (with >10,000 eliminated) for
each block.

COMPUTE
MeanBlock3=MEAN(upTC1,upTC2,upTC3,upTC4,upTC5,upTC6,upTC7,upTC8,
upTC9,upTC10,upTC11,upTC12,upTC13,upTC14,upTC15,upTC16,upTC17,up
TC18,upTC19,upTC20)+600.

VARIABLE LABELS MeanBlock3 'IAT mean block 3'.
EXECUTE.

COMPUTE
MeanBlock4=MEAN(upTC21,upTC22,upTC23,upTC24,upTC25,upTC26,upTC27
,upTC28,upTC29,upTC30,upTC31,upTC32,upTC33,upTC34,upTC35,upTC36,
upTC37,upTC38,upTC39,upTC40,upTC41,upTC42,upTC43,upTC44,upTC45,u

pTC46,upTC47,upTC48,upTC49,upTC50,upTC51,upTC52,upTC53,upTC54,up
TC55,upTC56,upTC57,upTC58,upTC59,upTC60)+600.

VARIABLE LABELS MeanBlock4 'IAT mean block 4'.
EXECUTE.

COMPUTE
MeanBlock6=MEAN(upTC61,upTC62,upTC63,upTC64,upTC65,upTC66,upTC67
,upTC68,upTC69,upTC70,upTC71,upTC72,upTC73,upTC74,upTC75,upTC76,
upTC77,upTC78,upTC79,upTC80)+600.

VARIABLE LABELS MeanBlock6 'IAT mean block 6'.
EXECUTE.

COMPUTE
MeanBlock7=MEAN(upTC81,upTC82,upTC83,upTC84,upTC85,upTC86,upTC87
,upTC88,upTC89,upTC90,upTC91,upTC92,upTC93,upTC94,upTC95,upTC96,
upTC97,upTC98,upTC99,upTC100,upTC101,upTC102,upTC103,upTC104,upT
C105,upTC106,upTC107,upTC108,upTC109,upTC110,upTC111,upTC112,upT
C113,upTC114,upTC115,upTC116,upTC117,upTC118,upTC119,upTC120)+60
0.

VARIABLE LABELS MeanBlock7 'IAT mean block 7'.
EXECUTE.

* Compute pooled SD for all trials in B3 & B6; another for B4 &
B7.

COMPUTE SD_Block3_6=SD(t1, t2, t3, t4, t5, t6, t7, t8, t9, t10,
t11, t12, t13, t14, t15, t16, t17, t18, t19, t20,t61, t62, t63,

t64, t65, t66, t67, t68, t69, t70, t71, t72, t73, t74, t75, t76, t77, t78, t79, t80).

VARIABLE LABELS SD_Block3_6 'IATsd_B3_6'.

EXECUTE.

COMPUTE
SD_Block4_7=sd(t21,t22,t23,t24,t25,t26,t27,t28,t29,t30,t31,t32,t33,t34,t35,t36,t37,t38,t39,t40,t41,t42,t43,t44,t45,t46,t47,t48,t49,t50,t51,t52,t53,t54,t55,t56,t57,t58,t59,t60,t81,t82,t83,t84,t85,t86,t87,t88,t89,t90,t91,t92,t93,t94,t95,t96,t97,t98,t99,t100,t101,t102,t103,t104,t105,t106,t107,t108,t109,t110,t111,t112,t113,t114,t115,t116,t117,t118,t119,t120).

VARIABLE LABELS SD_Block4_7 'IATsdB4_7'.

EXECUTE.

  * Eliminate trials of >10,000 from the original variable "t".

DO REPEAT
Time =t1 t2 t3 t4 t5 t6 t7 t8 t9 t10 t11 t12 t13 t14 t15 t16 t17 t18 t19 t20 t21 t22 t23 t24 t25 t26 t27 t28 t29 t30 t31 t32 t33 t34 t35 t36 t37 t38 t39 t40 t41 t42 t43 t44 t45 t46 t47 t48 t49 t50 t51 t52 t53 t54 t55 t56 t57 t58 t59 t60 t61 t62 t63 t64 t65 t66 t67 t68 t69 t70 t71 t72 t73 t74 t75 t76 t77 t78 t79 t80 t81 t82 t83 t84 t85 t86 t87 t88 t89 t90 t91 t92 t93 t94 t95 t96 t97 t98 t99 t100 t101 t102 t103 t104 t105 t106 t107 t108 t109 t110 t111 t112 t113 t114 t115 t116 t117 t118 t119 t120/Correct = c1 c2 c3 c4 c5 c6 c7 c8 c9 c10 c11 c12 c13 c14 c15 c16 c17 c18 c19 c20 c21 c22 c23 c24 c25 c26 c27 c28 c29 c30 c31 c32 c33 c34 c35 c36 c37 c38 c39 c40 c41 c42 c43 c44 c45 c46 c47 c48 c49 c50 c51 c52 c53 c54 c55 c56 c57 c58 c59 c60 c61 c62 c63 c64 c65 c66 c67

```
c68 c69 c70 c71 c72 c73 c74 c75 c76 c77 c78 c79 c80 c81 c82 c83
c84 c85 c86 c87 c88 c89 c90 c91 c92 c93 c94 c95 c96 c97 c98 c99
c100 c101 c102 c103 c104 c105 c106 c107 c108 c109 c110 c111 c112
c113 c114 c115 c116 c117 c118 c119 c120/TC = tcorrected1 to
tcorrected120.
if (Time >= 10000) Correct = 2.
end repeat.
EXECUTE.


* Replace each error latency with block mean + 600ms.


DO REPEAT
Time =t1 t2 t3 t4 t5 t6 t7 t8 t9 t10 t11 t12 t13 t14 t15 t16 t17
t18 t19 t20 /
Correct = c1 c2 c3 c4 c5 c6 c7 c8 c9 c10 c11 c12 c13 c14 c15 c16
c17 c18 c19 c20 /
TC = incorrectreplaced1 to incorrectreplaced20.
if (Correct = 1) TC = Time.
if (Correct=0) TC= (MeanBlock3 + 600).
if (Correct= 2) TC= 9999999.
end repeat.
EXECUTE.


DO REPEAT
Time = t21 t22 t23 t24 t25 t26 t27 t28 t29 t30 t31 t32 t33 t34
t35 t36 t37 t38 t39 t40 t41 t42 t43 t44 t45 t46 t47 t48 t49 t50
t51 t52 t53 t54 t55 t56 t57 t58 t59 t60 /
Correct = c21 c22 c23 c24 c25 c26 c27 c28 c29 c30 c31 c32 c33
c34 c35 c36 c37 c38 c39 c40 c41 c42 c43 c44 c45 c46 c47 c48 c49
c50 c51 c52 c53 c54 c55 c56 c57 c58 c59 c60  /
TC = incorrectreplaced21 to incorrectreplaced60.
```

```
if (Correct = 1) TC = Time.

if (Correct=0) TC= (MeanBlock4 +600).

if (Correct= 2) TC= 9999999.

end repeat.

EXECUTE.


DO REPEAT

Time =t61 t62 t63 t64 t65 t66 t67 t68 t69 t70 t71 t72 t73 t74
t75 t76 t77 t78 t79 t80 /

Correct = c61 c62 c63 c64 c65 c66 c67 c68 c69 c70 c71 c72 c73
c74 c75 c76 c77 c78 c79 c80 /

TC = incorrectreplaced61 to incorrectreplaced80.

if (Correct = 1) TC = Time.

if (Correct=0) TC = (MeanBlock6 +600).

if (Correct= 2) TC= 9999999.

end repeat.

EXECUTE.


DO REPEAT

Time = t81 t82 t83 t84 t85 t86 t87 t88 t89 t90 t91 t92 t93 t94
t95 t96 t97 t98 t99 t100 t101 t102 t103 t104 t105 t106 t107 t108
t109 t110 t111 t112 t113 t114 t115 t116 t117 t118 t119 t120
/Correct = c81 c82 c83 c84 c85 c86 c87 c88 c89 c90 c91 c92 c93
c94 c95 c96 c97 c98 c99 c100 c101 c102 c103 c104 c105 c106 c107
c108 c109 c110 c111 c112 c113 c114 c115 c116 c117 c118 c119 c120
/TC = incorrectreplaced81 to incorrectreplaced120.

if (Correct = 1) TC = Time.

if (Correct=0) TC = (MeanBlock7+600).

if (Correct= 2) TC= 9999999.

end repeat.
```

EXECUTE.

RECODE incorrectreplaced1 to incorrectreplaced120
(9999999=SYSMIS) (ELSE=Copy) INTO Allcorrected1 to
Allcorrected120.
EXECUTE.

COMPUTE
FinalMeanBlock3=MEAN(Allcorrected1,Allcorrected2,Allcorrected3,A
llcorrected4,Allcorrected5,Allcorrected6,Allcorrected7,Allcorrec
ted8,Allcorrected9,Allcorrected10,Allcorrected11,Allcorrected12,
Allcorrected13,Allcorrected14,Allcorrected15,Allcorrected16,Allc
orrected17,Allcorrected18,Allcorrected19,Allcorrected20).

VARIABLE LABELS FinalMeanBlock3 'IAT mean block 3 after
corrections'.
EXECUTE.

COMPUTE
FinalMeanBlock4=MEAN(Allcorrected21,Allcorrected22,Allcorrected2
3,Allcorrected24,Allcorrected25,Allcorrected26,Allcorrected27,Al
lcorrected28,Allcorrected29,Allcorrected30,Allcorrected31,Allcor
rected32,Allcorrected33,Allcorrected34,Allcorrected35,Allcorrect
ed36,Allcorrected37,Allcorrected38,Allcorrected39,Allcorrected40
,Allcorrected41,Allcorrected42,Allcorrected43,Allcorrected44,All
corrected45,Allcorrected46,Allcorrected47,Allcorrected48,Allcorr
ected49,Allcorrected50,Allcorrected51,Allcorrected52,Allcorrecte
d53,Allcorrected54,Allcorrected55,Allcorrected56,Allcorrected57,
Allcorrected58,Allcorrected59,Allcorrected60).

VARIABLE LABELS FinalMeanBlock4 'IAT mean block 4 after
corrections'.

```
EXECUTE.

COMPUTE
FinalMeanBlock6=MEAN(Allcorrected61,Allcorrected62,Allcorrected6
3,Allcorrected64,Allcorrected65,Allcorrected66,Allcorrected67,Al
lcorrected68,Allcorrected69,Allcorrected70,Allcorrected71,Allcor
rected72,Allcorrected73,Allcorrected74,Allcorrected75,Allcorrect
ed76,Allcorrected77,Allcorrected78,Allcorrected79,Allcorrected80
).

VARIABLE LABELS  FinalMeanBlock6 'IAT mean block 6 after
corrections'.
EXECUTE.

COMPUTE
FinalMeanBlock7=MEAN(Allcorrected81,Allcorrected82,Allcorrected8
3,Allcorrected84,Allcorrected85,Allcorrected86,Allcorrected87,

Allcorrected88,Allcorrected89,Allcorrected90,Allcorrected91,Allc
orrected92,Allcorrected93,Allcorrected94,Allcorrected95,Allcorre
cted96,Allcorrected97,

Allcorrected98,Allcorrected99,Allcorrected100,Allcorrected101,Al
lcorrected102,Allcorrected103,Allcorrected104,Allcorrected105,Al
lcorrected106,Allcorrected107,Allcorrected108,Allcorrected109,Al
lcorrected110,Allcorrected111,Allcorrected112,Allcorrected113,Al
lcorrected114,Allcorrected115,Allcorrected116,Allcorrected117,Al
lcorrected118,Allcorrected119,Allcorrected120).

VARIABLE LABELS  FinalMeanBlock7 'IAT mean block 7 after
corrections'.
EXECUTE.

COMPUTE DIFF_short=FinalMeanBlock6 - FinalMeanBlock3.
```

EXECUTE.

COMPUTE DIFF_long= FinalMeanBlock7 - FinalMeanBlock4.

EXECUTE.

COMPUTE quotshort=DIFF_short / SD_Block3_6.

EXECUTE.

COMPUTE quotlong=DIFF_long / SD_Block4_7.

EXECUTE.

COMPUTE D=MEAN(quotshort,quotlong).

EXECUTE.

DELETE VARIABLES tcorrected1 to allcorrected120.

execute.

*Old algorithm

USE ALL.

execute.

DATASET ACTIVATE DataSet1.

RECODE t23 t24 t25 t26 t27 t28 t29 t30 t31 t32 t33 t34 t35 t36
t37 t38 t39 t40 t41 t42 t43 t44 t45 t46 t47 t48 t49 t50 t51 t52
t53 t54 t55 t56 t57 t58 t59 t60 t83 t84 t85 t86 t87 t88 t89 t90
t91 t92 t93 t94 t95 t96 t97 t98 t99 t100 t101 t102 t103 t104
t105 t106 t107 t108 t109 t110 t111 t112 t113 t114 t115 t116 t117
t118 t119 t120 (Lowest thru 299.99=300) (10000.00001 thru

Highest=10000) (ELSE=Copy) INTO oldt23 oldt24 oldt25 oldt26 oldt27 oldt28 oldt29 oldt30 oldt31 oldt32 oldt33 oldt34 oldt35 oldt36 oldt37 oldt38 oldt39 oldt40 oldt41 oldt42 oldt43 oldt44 oldt45 oldt46 oldt47 oldt48 oldt49 oldt50 oldt51 oldt52 oldt53 oldt54 oldt55 oldt56 oldt57 oldt58 oldt59 oldt60 oldt83 oldt84 oldt85 oldt86 oldt87 oldt88 oldt89 oldt90 oldt91 oldt92 oldt93 oldt94 oldt95 oldt96 oldt97 oldt98 oldt99 oldt100 oldt101 oldt102 oldt103 oldt104 oldt105 oldt106 oldt107 oldt108 oldt109 oldt110 oldt111 oldt112 oldt113 oldt114 oldt115 oldt116 oldt117 oldt118 oldt119 oldt120.

EXECUTE.

do repeat

x= oldt23 oldt24 oldt25 oldt26 oldt27 oldt28 oldt29 oldt30 oldt31 oldt32 oldt33 oldt34 oldt35 oldt36 oldt37 oldt38 oldt39 oldt40 oldt41 oldt42 oldt43 oldt44 oldt45 oldt46 oldt47 oldt48 oldt49 oldt50 oldt51 oldt52 oldt53 oldt54 oldt55 oldt56 oldt57 oldt58 oldt59 oldt60 oldt83 oldt84 oldt85 oldt86 oldt87 oldt88 oldt89 oldt90 oldt91 oldt92 oldt93 oldt94 oldt95 oldt96 oldt97 oldt98 oldt99 oldt100 oldt101 oldt102 oldt103 oldt104 oldt105 oldt106 oldt107 oldt108 oldt109 oldt110 oldt111 oldt112 oldt113 oldt114 oldt115 oldt116 oldt117 oldt118 oldt119 oldt120/y= oldlog1 to oldlog76.

compute y=ln(x).

end repeat.

Execute.

compute compblock= mean(oldlog1 to oldlog38).

compute incompblock= mean (oldlog39 to oldlog76).

compute OldD= incompblock - compblock.

execute.

```
delete variables oldt23 to oldlog76.


EXECUTE.
```