Lund Psychological Reports

Volume 10, No. 1 2009

The Indirect and Direct Attitude Measure (IDAM)

Jean-Christophe Rohner, Johan Mårtensson, Martin Geisler, Samantha Sinclair, & Josefin Zetterberg

Department of Psychology

LUND UNIVERSITY

Lund Psychological Reports

Editor: Magnus Lindgren
ISSN 1404-8035

Abstract

In this paper we present and evaluate a new method to assess implicit attitudes as associations that are introspectively unidentified traces of past experience (Greenwald & Banaji, 1995): The indirect and direct attitude measure (IDAM). By using recognition-bias in a recognition memory task, we can measure the conditional probability of positive evidence for an association on an indirect test provided that there is positive evidence for the absence of the same association on a direct test; i.e. if an attitude is truly implicit. Five experiments, with a total of 246 participants, revealed that IDAM could measure truly implicit attitudes towards a variety of stimuli (such as affective images and attractive faces). Effect sizes where highly variable, however, and in a certain condition the IDAM was ineffective. Taken together, the research indicated that the IDAM is a valid measure of implicit attitude, but that more research is needed to ascertain if IDAM can be used to predict behaviour.

Introduction

There are several reasons for wanting to obtain valid and reliable attitude measurements. Such measurements can yield estimates of public opinion, allow prediction of behaviour and be used to evaluate the effectiveness of various attitude-change interventions. In this paper we present a new technique to measure implicit attitudes, IDAM. In five experiments we evaluate IDAMs validity, the effects of different procedural variations, as well as different algorithms to calculate implicit attitudes from IDAM responses.

There are various formal definitions of attitudes, but a common denominator is that attitudes are viewed as mental representations that link a social concept to either negative or positive valence; attitudes are thought to influence cognition, emotion and behaviour (Eagly & Chaiken, 1998). The attitude concept is often divided into explicit and implicit components. Explicit attitudes are described as mental representations that can be made available to awareness and therefore reported by means of introspection. Implicit attitudes are “introspectively unidentified (or inaccurately identified) traces of past experience that mediate favourable or unfavourable feeling, thought or action” p 8 (Greenwald & Banaji, 1995).

The term implicit attitude is relatively new. During the early history of social psychology it was not uncommon to simply assume that people are aware of their preferences and that they are able to report this information on the basis of simple introspection; the apparently widespread use of questionnaire techniques supports this conclusion (e.g. Greenwald, 1990 and Banaji & Greenwald, 1994). Soon, however, researchers realized that self-report techniques are easily distorted by demand effects (e.g. Orne, 1962) and self-presentation strategies (e.g. Tedeschi, Schlenker, Bonoma, 1971). I.e. people were reporting the kind of attitudes that they believed the researcher wanted them to report and/or the kind of attitudes that they believed were socially desirable. Nisbett and Wilson (1977) convincingly demonstrated the many problems that are connected with self-report techniques.

During the nineties the popularity of various indirect measurement procedures to estimate implicit attitudes and stereotypes has increased dramatically (Greenwald & Banaji, 1995). A keyword search in psych info for “attitude and implicit” conducted in 2008 gave 1425 hits for example. In a review, Fazio and Olson (2003) write “Over the last few years, there has been a surge of interest in the use of implicit measurement techniques in social psychological research” p 2.

Indirect measurement procedures are used as an indicator of implicit attitudes. These measures usually have two things in common: (1) they provide a quantitative estimate of the strength of association between a concept (e.g. female) and an attribute (e.g. positive), and (2) they do not rely on asking subjects if they think that the concept is characterized by the attribute. I.e. it is possible to measure somebody’s attitude without asking them about their attitude. Indirect measures can be compared to direct measures which use self-reports as an indicator of someone’s attitude. Therefore, the difference between indirect and direct procedures lies in whether a participant is aware of the purpose of a measure (or what is indicated by a certain response) or not. If a participant believes that a certain response is an indicator of A but the response is used as an indicator of B, the response can be said to be an indirect measure of B and a direct measure of A (see Greenwald & Banaji, 1995).

The indirect measures that have been used in social psychology have often been inspired by advances in cognitive psychology. Semantic priming can serve as an example: When a person has to categorize a target stimulus (e.g. judge whether *ambulance* is a word or not), the categorization can be made faster if a semantically related prime word (e.g. *hospital*) is presented immediately before (e.g. Neely, 1977). The categorization is made faster because the person possesses a mental representation in which the two concepts are associated. When researchers used this technique to measure attitudes in social psychological research, they found that Caucasian Americans categorized positive target words faster when they were preceded by an image of a Caucasian American than when they were preceded by an image of an African American. They concluded that Caucasian Americans have more positive attitudes towards Caucasian Americans than towards African Americans (Gartner & McLaughlin, 1983; Dovidio, Evans & Tyler, 1986). Other studies also revealed that the prime affected target categorizations independently of the participant's intentions: target categorizations (valence words) were faster even though the interval between prime (faces of different ethnic groups) was as short as 250 ms (Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Bargh, Chaiken, Govender, & Pratto, 1992). All in all automatic attitude activation appears to be a well established phenomenon (Wegner & Bargh, 1998).

Another very common indirect measure of attitudes is the implicit association test (IAT; Greenwald, McGee, & Schwartz, 1998). In this technique participants are required to categorize stimuli that represent one of two social concepts (e.g. *George Bush* and *Barack Obama*) according to what concept they belong to (e.g. *republican* and *democrat*) and to categorize stimuli that represent positive and negative valence (e.g. *war* and *peace*) according to valence (e.g. *negative* and *positive*). Concept stimuli and valence stimuli are intermixed within the same test so that a concept stimulus may appear on one trial and a valence stimulus may appear on the next trial. The same response is used for a certain concept categorization and for a certain valence categorization; e. g. during one part of the test participants may be required to use one response button for both *democrat* and *positive*, and required to use another response button for both *republican* and *negative*; during another part of the test participants may be required to use one response button for both *democrat* and *negative* and another response button for both *republican* and *positive*. When a certain response is mapped to two things that a person associates the response is made faster than when the response is mapped to two things that a person does not associate.

Techniques that involve priming and variants of the IAT are the most common but there are also other indirect measures of attitudes and stereotypes. Fazio and Olson (2003) list at least 7 techniques, in addition to priming and the IAT. These include procedures that involve such things as word stem completion, analyses of language content, facial EMG, fMRI and EEG.

In some cases, the content of an implicit attitude may disagree with the outcome on a direct test of the attitude; e.g. participants may show evidence of prejudice on an implicit measure but not report prejudice on an explicit measure. This pattern is important because it suggests that implicit attitudes are distinct from explicit attitudes. Greenwald and Banaji (1995) write “Possibly, the evaluative content of this implicit attitude may disagree with results from a direct measure of attitude toward B; such disagreement, referred to as a dissociation of implicit and explicit attitudes, is especially interesting and perhaps most dramatically indicates the value of the implicit attitude construct.” p 8.

A number of research projects have studied relations between direct and indirect attitude measures. This question is important because non-existent or negative correlations between indirect and direct tests would support the conclusion that implicit attitudes exist and that they are distinct from explicit attitudes; i.e. that humans have associations between certain concepts and certain attributes that are unrelated to their conscious beliefs about these associations. The definition of implicit attitudes that is provided by Greenwald and Banaji (1995) implies that humans have mental representations containing links between social concepts and valence attributes that they are unaware of. If this definition is correct humans should (1) show that they associate a certain concept with negative or positive valence in an indirect measure of attitudes, (2) show a lack of awareness about the association that was evident on the indirect measure. So the definition would be supported if indirect attitude measures were non correlated or negatively correlated with direct attitude measures. When this hypothesis is tested on empirical data, in fact, relations between indirect and direct measures of attitudes tend to be low or non existent (Blair, 2001; Dovidio, Kawakami, Beach, 2001; Brauer, Wasel, Niedenthal, 2000). One interpretation of this result is that implicit attitudes exist according to the definition given by Greenwald and Banaji (1995).

Fazio and Olson (2003) on the other hand, suggest that low correlations between indirect and direct attitude measures can be explained by a moderating effect of motivation to control prejudice, and not by the fact that participants are unaware of attitudes that are measured indirectly. Attitudes are activated automatically when relevant social objects are encountered; attitude expression is, however, under the influence of controlled processes. People are motivated to present themselves in a positive way. When they have the ability to control the expression of socially unacceptable attitudes and when they are motivated to do so, they will modify the expression so that it becomes more socially acceptable (Fazio & Olson, 2003). Direct and indirect attitude measures should thus only correlate when people are not motivated to alter how they express an attitude. Empirical evidence seems to support this hypothesis. In a meta-analysis Nosek (2005) found that the relation between indirect and direct measures was moderated by self presentation concerns. However, the moderating effect of self presentation was quite weak and could not fully account for the relation between indirect and direct tests.

There seem to be three problems with the research that has examined relations between indirect and direct attitudes: (1) definitions for direct and indirect measurement procedures cannot be proven to be logically independent; (2) the stimuli and responses that have been used in indirect and direct attitude measures are very different; (3) evidence for an association on indirect measures has not been demonstrated to be contingent on evidence for the absence of the association on direct measures.

Problems with typical measures of implicit attitudes

Implicit attitudes exist (see Greenwald and Banaji, 1995) if participants show evidence of an association on an indirect measure (IM) but no evidence of an association on a direct measure (DM). I.e. scores on indirect and direct test should be non-correlated or negatively correlated. This question is meaningful only if evidence can be provided that the definitions of indirect and direct tests are logically independent. Assume that IM has the characteristics A and B and that DM has the characteristic A but not B (A and B could be different questions on a questionnaire for example), so we could write IM = A + B and DM = A. In this case, DM and IM are *always* going to be correlated, no matter what values A and B have. Fazio and Olson (2003) argued that typical indirect attitude measures, like the IAT, often also have a conscious component; no such measure is completely free of “contamination” from consciousness. In their words, “Nothing about our current implicit measurement procedures, be it a priming method, the IAT, or one of the other techniques mentioned earlier, guarantees that participants are unaware of their attitudes” p 9. This is analogous to saying that IAT = implicit component + explicit component and, assuming that typical questionnaire measures tap explicit attitudes, questionnaire = explicit component. If this is the case, an IAT and a questionnaire will always correlate, no matter what values people have on the implicit and explicit components. The Greenwald and Banaji (1995) definition could never be true. IDAM has the advantage that its indirect and direct attitude measures can be proven to be orthogonal (i.e. logically independent).

The second problem concerns the fact that typical indirect and direct measurement techniques usually have very different formats. Indirect measures are often based on some sort of reaction-time measurement; direct measures are usually based on self reports or ratings. Moreover; indirect tests often involve simple word stimuli or image stimuli, while direct tests sometimes involve complex sentences. If low correlations are found between indirect and direct tests of attitudes this could be due to the fact that the two types of procedures have different formats, and not because participants are unaware of the associations that are revealed in indirect measures. Nosek (2005) makes a similar argument. If a measurement procedure used similar responses and stimuli for direct and indirect tests, we might find that indirect and direct tests are correlated; such a result would be incongruent with the Greenwald and Banaji (1995) definition of implicit attitudes. If, on the other hand, such a procedure revealed low or non-existent correlation between direct and indirect tests, this would yield strong support for the Greenwald and Banaji (1995) definition. The present technique provides direct and indirect measurement techniques that involve exactly the same responses and exactly the same stimuli.

Greenwald and Banaji (1995) define implicit attitudes as “introspectively unidentified (or inaccurately identified) traces of past experience that mediate favourable or unfavourable feeling, thought or action" p 8 (Greenwald & Banaji, 1995). Typical indirect attitude measures that are used to indicate implicit attitude do, however, not show that evidence of an association on an indirect test is *contingent* on there being evidence that the same association is introspectively unidentified or inaccurately identified on a direct measure. According to Fazio and Olson (2003) research on implicit attitudes almost never examine if an implicit measure indicates a construct of which participants are unaware. In a review of implicit measures of attitudes, De Houwer and Moors (2007) write "To summarize, despite the common claim that implicit measures are unconscious measures, these claims are most often ill specified and lack supporting evidence. More research on this issue is urgently needed." p 191. A measurement procedure that is a valid indicator of implicit attitude according to the Greenwald and Banaji (1995) - definition, should have the following format: p (positive evidence for an association on an indirect measure | positive evidence for the absence of the association on a direct measure); i.e. the conditional probability of positive evidence for an association on the indirect measure, given that the direct measure shows that the same association is absent. The IDAM yields such an estimate.

The Indirect and Direct Attitude Measure (IDAM)

Theoretical and empirical rationale

Several studies have shown that episodic long term memory interacts with semantic long term memory when people try to recall an episode, so that information about what actually happened in the episode is confounded with knowledge and expectations about the episode (Bartlett, 1932; Roediger, Watson, McDermott, & Gallo, 2001; Roediger & McDermott, 1995; Schacter, 2001). A simple and very reliable demonstration of this phenomenon is that people think that they have seen a word that represents a concept (e.g. *hospital*) in a memory test when words that represent instances of the concept are presented during an earlier study phase (e.g. *doctor, nurse, medicine*), even if the concept word was never shown. The confounding of episodic and semantic memory leads to a recognition bias where both old (studied) and new (not studied) items attract more "old" responses. A plausible explanation of this phenomenon is that stimuli that are conceptually linked are processed more fluently; fluency, in turn, is interpreted as a sign of conscious memory (i.e. fluency is misattributed to the stimuli appearing at study) (Kelley & Rhodes, 1990). It is understandable that people make this attribution, because stimuli that have been encoded are in fact are processed more easily (e.g. Roediger, 1990).

In the IDAM, recognition bias from an explicit memory test is used as an indirect measure of attitude. Attitudes are defined as mental representations that connect a concept to an attribute; in essence they are semantic memory representations. Semantic memory representations have been shown to produce recognition biases for non social stimuli; by being semantic memory representations attitudes should also produce recognition biases. This deduction is also supported by a large number of empirical studies. Two meta-analyses show that attitude congruent information produces more recognition bias than attitude incongruent information, without necessarily producing more accurate memory than incongruent information (Stangor & McMillan, 1992; Eagly, Chen, Chaiken, & Shaw-Barnes, 1999).

The measurement procedure

IDAM can be used to obtain independent direct, indirect and implicit measures of attitudes towards ideologies, artefacts, humans and behaviours. The technique uses an explicit recognition memory test involving combinations of stimuli that represent concepts and stimuli that represent attributes as the indirect attitude measure. Several types of concept stimuli can be used: words, names, images, sentences, etc. At study a number of stimuli that represent concept + attribute combinations are shown together on a screen (e.g. Anna + intelligent, Sara + unintelligent, John + smart, Paul + dumb, for the concepts female and male, and the attributes positive and negative). At test, study combinations are presented with an equal number of new combinations that are created by repairing study combinations (e.g. Anna + smart, Sara + dumb, John + intelligent, Paul + unintelligent). The study phase only contains the instances that were presented at test (i.e. no new stimuli are shown at test, only new combinations). New combinations also preserve the mapping between concept category instances and the attribute category that existed at study (i.e. if Anna + intelligent was an old combination, Anna + smart could be a new combination, but Anna + unintelligent would not be a new combination). The attribute category that is presented together with a certain instance of the concept category is thus the same for both old and new combinations. Table 1 shows a general representation of how the test is carried out (with the concept categories female and male and the attribute categories negative and positive as examples).

Table 1.
Conditions in IDAM

Study
Concept A: Instance A1 (Anna) + Attribute L: Instance L1 (intelligent)
Concept A: Instance A2 (Sara) + Attribute M: Instance M1 (unintelligent)
Concept B: Instance B1 (John) + Attribute L: Instance L2 (smart)
Concept B: Instance B2 (Paul) + Attribute M: Instance M2 (dumb)

Test - Old items
Concept A: Instance A1 (Anna) + Attribute L: Instance L1 (intelligent)
Concept A: Instance A2 (Sara) + Attribute M: Instance M1 (unintelligent)

Concept B: Instance B1 (John) + Attribute L: Instance L2 (smart)
Concept B: Instance B2 (Paul) + Attribute M: Instance M2 (dumb)
Test - New items
Concept A: Instance A1 (Anna) + Attribute L: Instance L2 (smart)
Concept A: Instance A2 (Sara) + Attribute M: Instance M2 (dumb)
Concept B: Instance B1 (John) + Attribute L: Instance L1 (intelligent)
Concept B: Instance B2 (Paul) + Attribute M: Instance M1 (unintelligent)

Typically we use at least 20 instances for each concept category and attribute category combination in order to obtain reliable measurements.

Instances of a given attribute category (e.g. intelligent and smart) are randomly paired with instances that represent a given concept category (e.g. Anna and Sara), and this assignment is made anew for each participant. Similarly, each participant receives a new random order of conditions at both study and test. Direct, indirect and implicit attitude measures are then obtained for each combination of concept categories and attribute categories, i.e. for A + L, A + M, B + L, and B + M (in the example female + positive, female + negative, male + positive, male + negative).

During the test phase the task is to judge if each combination was shown at study or not and to judge if one agrees or not with the information presented in each combination. One of four responses alternatives can be chosen: 1. Old & Agree, 2. Old & Disagree, 3. New & Agree, 4. New & Disagree. Judging that a combination was old but disagreeing with the information in the combination would produce the response Old & Disagree for example. The four response alternatives are presented on a screen together with the concept instances and attribute instances and their position varies randomly from trial to trial. Participants always respond by clicking with the computer mouse and there is no time-limit in this task. Observed frequencies in each condition (A + L, A + M, B + L, and B + M) can then be calculated separately for old and new combinations as shown in table 2.

Table 2.

Response categories for the IDAM

Response	Item type	Meaning	Variable name
Old & Agree	Old	Hit & Agree	HA
Old & Agree	New	False alarm & Agree	FA
Old & Disagree	Old	Hit & Disagree	HD
Old & Disagree	New	False alarm & Disagree	FD
New & Agree	Old	Miss & Agree	MA

New & Agree	New	Correct rejection & Agree	CA
New & Disagree	Old	Miss & Disagree	MD
New & Disagree	New	Correct rejection & Disagree	CD

Operational definitions

Direct attitude measure

For the direct measure of attitude, DM, we simply use the frequency of “agree” - responses relative to “disagree” - responses, as shown in equation 1.

$$DM = HA + FA - HD - FD + MA + CA - MD - CD \quad (1)$$

Indirect attitude measures

For the indirect measure of attitude, IM, we try and evaluate two different algorithms. One represents the difference between the frequency of “old”-responses relative to the frequency of “new”-responses, as shown in equation 2 (L for linear contrast). I.e. we postulate that IM increases linearly when “old” responses increase linearly and when “new” responses decrease linearly.

$$IM_L = HA + FA + HD + FD - MA - CA - MD - CD \quad (2)$$

Note that the frequency of “agree” - responses and the frequency of “disagree” - responses, per se, do not affect the magnitude of IM_L in equation 2. A very important property of this measure of IM is also that the operationalizations of DM and IM_L are orthogonal; i.e. we can show that DM and IM_L are independent. In other words, in the space of HA, FA, HD, FD, MA, CA, MD and CD, DM and IM_L are perpendicular axes.

The other algorithm is based on multinomial modelling and is similar to the approach used in the Two High Threshold model for recognition memory (see Snodgrass & Corwin, 1988). In this operationalization we make the following assumptions:

- A bias to respond “old” is an indirect indicator that a person has a certain attitude (IM).
- A bias to respond “new” is an indirect indicator that a person does not have a certain attitude (1 – IM)
- Old items are recognized as old with the same probability that new items are recognized as new ($R_{OLD} = R_{NEW} = R$).

- Hits (HA or HD) occur when an item is recognized as old (R), or when the item is not recognized as old (1 – R) and a participant has a bias to respond “old” (IM), (see equations 3 and 5).
- False alarms (FA or FD) occur when an item is not recognized as new (1 – R) and a participant has a bias to respond “old” (IM) (see equations 4 and 6).
- Correct rejections (CA or CD) occur when an item is recognized as new (R), or when an item is not recognized as new (1 – R) and a participant has a bias to respond “new” (1 – IM) (see equations 7 and 9).
- Misses (MA or MD) occur when an item is not recognized as old (1 – R) and a participant has a bias to respond “new” (1 – IM) (see equations 8 and 10).

$$HA = R + (1 - R)\,IM \qquad (3)$$
$$FA = (1 - R)\,IM \qquad (4)$$
$$HD = R + (1 - R)\,IM \qquad (5)$$
$$FD = (1 - R)\,IM \qquad (6)$$
$$CA = R + (1 - R)(1 - IM) \qquad (7)$$
$$MA = (1 - R)(1 - IM) \qquad (8)$$
$$CD = R + (1 - R)(1 - IM) \qquad (9)$$
$$MD = (1 - R)(1 - IM) \qquad (10)$$

With 2 unknowns (IM and R) and 8 equations, solving for IM (and R), we can get 4 estimates of IM. From equations 3 and 4 we get $IM = - FA / (- 1 - FA + HA)$, from equations 5 and 6 we get $IM = - FD / (- 1 - FD + HD)$, from equations 7 and 8 we get $IM = (- 1 + CA) / (-1 + CA - MA)$ and from equations 9 and 10 we get $IM = (- 1 + CD) / (- 1 + CD - MD)$. The indirect attitude measure is the combination of these four estimates, as shown in equation 11 (MN stands for “multinomial”).

$$IM_{MN} = - FA / (- 1 - FA + HA) - FD / (- 1 - FD + HD) + (- 1 + CA) / (-1 + CA - MA) + (- 1 + CD) / (- 1 + CD - MD) \qquad (11)$$

Note that, as with the other indirect attitude measure, the frequency of “agree” - and “disagree” - responses, per se, does not affect the magnitude of IM_{MN}. Equations 2 and 11 represent indirect measures of attitude towards a concept, independently of whether the direct measure shows the same attitude or not; i.e. IM_L | (Agree or Disagree) and IM_{MN} | (Agree or Disagree).

Implicit indirect attitude measures and explicit indirect attitude measures

The IDAM yields measures of implicit attitudes that are congruent with the definition “introspectively unidentified (or inaccurately identified) traces of past experience that mediate favourable or unfavourable feeling, thought or action” p 8 (Greenwald & Banaji, 1995). I.e. that implicit attitude is the conditional probability of positive evidence for an association on an indirect test, provided that there is positive evidence for the absence of the association on the direct test; see equation 12 and 13 (IIM stands for implicit indirect measure, L stands for linear, MN stands for multinomial). Compare this to the conditional probability of positive evidence for an association on the indirect measure, provided that there is positive evidence for the association on the direct measure; see equations 14 and 15 (EIM stands for explicit indirect measure).

$$IIM_L = HD + FD - MD - CD \quad (12)$$

$$IIM_{MN} = - FD / (- 1 - FD + HD) + (- 1 + CD) / (- 1 + CD - MD) \quad (13)$$

$$EIM_L = HA + FA - MA - CA \quad (14)$$

$$EIM_{MN} = - FA / (- 1 - FA + HA) + (- 1 + CA) / (-1 + CA - MA) \quad (15)$$

Overview of the present research

In this paper we examine the construct validity and various measurement properties of the IDAM. To study the IDAMs construct validity as an indicator of implicit attitude, we use concept categories for which other studies have demonstrated negative and positive attitudes: Images of physically attractive and physically unattractive individuals and pictures with emotional content from the International Affective Picture System (IAPS, Lang, Bradley, & Cuthbert, 2005). The attribute categories are *negative* and *positive*, with various words as instances. If the IDAM is a valid measure of implicit attitude we should find that IIM_L and/or IIM_{MN} have higher values for the concept category + attribute category combinations that have been shown to be congruent with participants' attitudes in other studies. We also study whether indirect attitude measures vary as a function of direct attitude measures. If indirect and direct measures are unrelated, we should be able to state with confidence that implicit attitudes are distinct from explicit attitudes. I.e. if indirect attitude measures were completely contingent on direct measurements (so that a participant only showed an attitude on the indirect test when the same attitude was indicated on the direct test), the importance of the implicit attitude construct would be considerably weakened.

We also study a number of methodological properties of the IDAM: The effect of using different algorithms to compute IIM and EIM, the effect of type of task at study and the effect of using a large or small number of concept category instances and attribute category instances. Ideally, the measure should be insensitive to different algorithms to compute IIM and insensitive to different procedural variations.

Experiment 1

In our first experiment we wanted to examine the IDAMs construct validity by using two classic social concepts: attractive and unattractive. A number of meta-analytic studies have shown that physically attractive targets are evaluated and treated more favourably than physically unattractive targets (Adams, 1982; Dion, Berscheid, & Walster, 1972; Eagly, Ashmore, Makhijani, & Longo, 1991; Feingold, 1992; Langlois et al., 2000). If the IDAM is able to measure implicit attitude we should find that IIM is higher for the combinations attractive + positive and unattractive + negative, than for the combinations attractive + negative and unattractive + positive.

Method

Participants

Thirty two student subjects in Sweden participated in experiment 1. Mean age was 25.38 (SD = 7.09). There were 69% female and 31% male participants in the sample.

Stimuli

A total of 142 greyscale images of faces were initially collected from Ekman and Friesen's Pictures of Facial Affect (1976), from the Karolinska Directed Emotional Faces (Lundkvist, Flykt, & Öhman, 1998) and from the Internet. Six judges (three female and three male) then rated these images for attractiveness on a four point scale that ranged from "very unattractive" to "very attractive". Forty faces that received the lowest and the highest average attractiveness ratings were used as attractive faces (20) and unattractive faces (20), respectively. An ANOVA, with stimuli as cases (and with the average across judges as the dependent variable), also confirmed that attractive faces received significantly higher average attractiveness ratings than unattractive faces, F (1, 38) = 231.14, $p < 0.05$, $\eta^2 = 0.86$, M = 3.22 and M = 1.54, respectively. Unattractive and attractive faces were matched on emotional expression (17 neutral and 3 happy in each category) and sex (10 female and 10 male in each category).

Six other judges (three female and three male), rated the valence ("positive" or "negative") of a total of 93 words that described personality traits, abilities and various psychological states. Twenty words that all judges rated as positive were used as positive words and 20 words that all judges rated as negative were used as negative words. Positive and negative words were matched on frequency of occurrence in Swedish newspapers and on word length. All words are listed in the appendix.

Procedure

Participants were tested individually or in pairs and they went through the IDAM according to the description in the introduction. Instances for the concept category *attractive* were 20 attractive faces, instances for the concept category *unattractive* were 20 unattractive faces, instances for the attribute category *positive* were 20 positive words, and instances for the attribute category *negative* were 20 negative words.

At study, participants performed a relational encoding task where they had to rate if a face + word combination on a given trial was similar to the face + word combination that had been shown on the previous trial. Ratings were made by clicking with the computer mouse on an alternative on a 4-point scale that ranged from “very dissimilar” to “very similar”. The purpose of this task was to focus attention on the faces and words. There was no time limit in this task. The study phase consisted of 40 trials in total. Participants then went through a distraction phase. Here they completed a vocabulary test in which they had to determine the meaning of each of 7 words by choosing one of 5 alternative answers. Words and response - alternatives were presented on the computer monitor (to the left and right, respectively) and participants made their choice by clicking on an answer with the computer-mouse. The test phase consisted of 80 trials.

Results

As a first step we calculated response frequencies for congruent and incongruent combinations. Congruent combinations consisted of *attractive* + *positive* and *unattractive* + *negative*. Incongruent combinations consisted of *attractive* + *negative* and *unattractive* + *positive*. Descriptive statistics for these frequencies are shown in table 3.

Table 3.

Response frequencies

	M		*SD*	
	Congruent	Incongruent	Congruent	Incongruent
HA	0.31	0.23	0.12	0.11
FA	0.14	0.09	0.08	0.10
CA	0.38	0.22	0.14	0.10
MA	0.26	0.11	0.12	0.07
HD	0.21	0.32	0.12	0.11
FD	0.13	0.17	0.10	0.12
CD	0.35	0.53	0.10	0.16
MD	0.22	0.34	0.11	0.13

Next, we calculated the IDAM - variables for congruent and incongruent combinations as described in the introduction: DM, IIM_L, EIM_L, IIM_{MN}, and EIM_{MN}. Before entering the statistical analysis each variable was inspected for outliers, defined as values with a two-tailed probability lower than 0.01 in a normal distribution. Outliers were replaced with the raw scores that correspond to 0.01/2 and 1 – 0.01/2, using an iterative procedure that only ended when all outliers were gone (i.e. by recalculating means, standard deviations and probabilities after each group of outliers had been removed); see Tabachnick and Fidell (2001). The variable that had the highest number of outliers had 6.25% outliers.

The indirect measures derived from the linear algorithm where then used as the dependent variable in a repeated measures ANOVA with congruence (2: congruent, incongruent) and measure (2: IIM_L, EIM_L) as within subjects factors. As shown in table 4 and figure 1 there was a significant congruence x measure interaction. Congruent combinations yielded higher values on IIM_L than incongruent combinations; EIM_L revealed the opposite pattern, with higher values on incongruent than congruent combinations (see table 4 and figure 1). When using the linear algorithm, the implicit indirect measure (IIM) thus could discriminate between congruent and incongruent combinations in the predicted direction. Moreover, implicit indirect attitudes were dissociated from explicit indirect attitudes.

The ANOVA that used measures from the multinomial algorithm as a dependent variable, more or less showed the same pattern of results, with a significant interaction between congruence and measure, and a significant effect of congruence on IIM_{MN} in the predicted direction (see table 4 and figure 1). The effect of congruence on IIM was, however, larger when using the multinomial algorithm (31% explained variance) than the linear algorithm (17% explained variance).

Table 4.

Results on the indirect measures

Algorithm	Effect	*F*	*p*	η^2
Linear	Congruence x Measure	17.86	< 0.05	0.37
	Congruence	0.61	0.44	0.02
	Measure	30.33	< 0.05	0.50
	Congruence | IIM_L	6.43	< 0.05	0.17
	Congruence | EIM_L	20.90	< 0.05	0.40
Multinomial	Congruence x Measure	21.35	< 0.05	0.41
	Congruence	0.15	0.71	0.01
	Measure	16.97	< 0.05	0.35
	Congruence | IIM_{MN}	14.22	< 0.05	0.31
	Congruence | EIM_{MN}	14.84	< 0.05	0.32

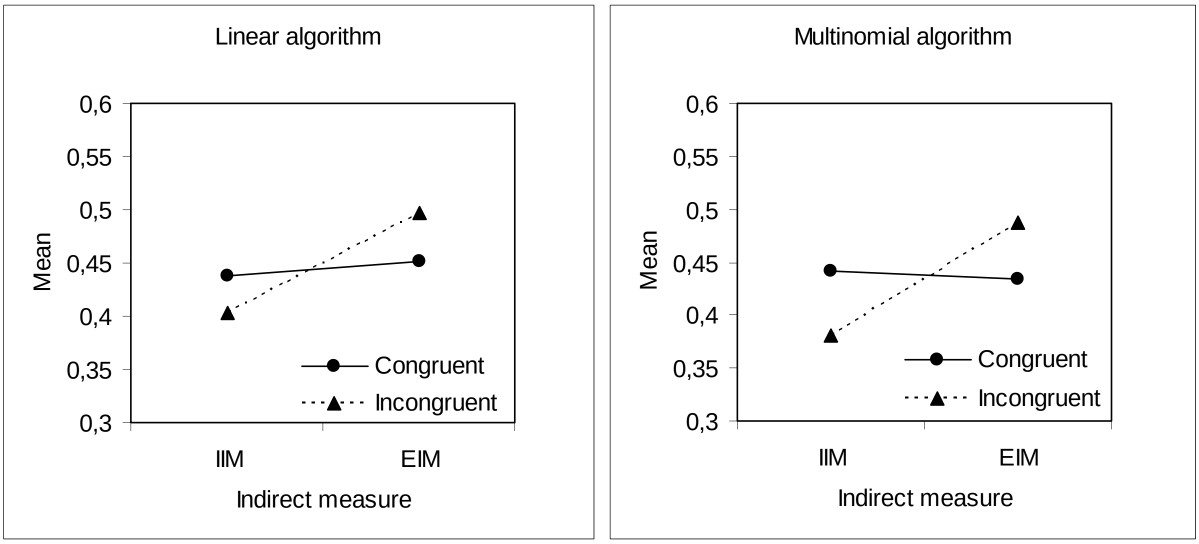

Figure 1. Averages for experimental conditions on the indirect measures

On the direct measure (DM) there was a strong main effect of congruence with an advantage for congruent combinations, F (1, 31) = 87.99, $p < 0.05$, $\eta^2 = 0.74$ (congruent combinations had a mean of 0.52 and a standard deviation of 0.05, incongruent combinations had a mean of 0.41 with a standard deviation of 0.05).

Discussion

The implicit indirect measure of the IDAM could discriminate between congruent combinations (attractive + positive and unattractive + negative) and incongruent combinations (attractive + negative and unattractive + positive). Using the multinomial algorithm we could predict about 30% of the variance in IIM_{MN} from the type of concept category + attribute category combination. This result supports the IDAMs construct validity as a measure of implicit attitude. The linear algorithm and the multinomial algorithm both yielded similar results which is reassuring.

Congruent combinations produced higher values than incongruent combinations on the implicit indirect measures (IIM) but incongruent combinations produced higher values than congruent combinations on the explicit indirect measure. This is an interesting result because it shows that implicit (indirect) and explicit (indirect) attitudes are dissociated; i.e. there is a reason to assume that implicit attitudes are separate from explicit attitudes. An explanation of this finding is that congruent combinations produce more fluency than incongruent combinations, and that the fluency that is produced by congruent combinations is only attributed to “old” when a person responds “disagree”. When a person responds “agree”, the fluency for congruent combinations can instead be attributed to “agree”. I.e. the only data point that increases due to fluency is IIM for congruent combinations thus producing the interaction. This explanation is similar to the explanation that has been offered for why mere exposure effects disappear sometimes. Bornstein (1992) observed that mere exposure effects decreased when participants remembered stimuli that were shown at study. I.e. when an item is remembered, fluency cannot be attributed to liking, fluency has instead to be attributed to the item being old.

Experiment 2

Experiment 1 employed attractive and unattractive faces as concept categories to validate IDAM as an implicit indirect measure of attitude. The purpose of experiment 2 was to study whether the implicit indirect attitude measure of IDAM could discriminate between a broader range of stimuli. Experiment 1 only employed two concept categories; in experiment 2 we wanted to know whether the results of experiment 1 would generalize across 10 additional stimulus categories. Experiment 2 used standardized stimuli from the International Affective Picture System as instances (IAPS, Lang, Bradley, & Cuthbert, 2005); the concept categories were: (1) negative and neutral actions, (2) negative and neutral animals, (3) negative and neutral environmental scenes, (4) negative and neutral humans, (5) negative and neutral objects. The attribute categories were negative and neutral, with words as specific instances. Congruent combinations consisted of negative images + negative words, and of neutral images + neutral words. Incongruent combinations consisted of negative images + neutral words, and neutral images + negative words. If the IDAM is a valid indirect implicit measure of attitude, we should find that its implicit indirect measures IIM_L and IIM_{MN} have higher values for congruent than for incongruent combinations. Moreover, the advantage for congruent combinations should be evident across image types (i.e. actions, animals, environmental scenes, humans and objects).

In experiment 2 we also wanted to evaluate a procedural variable. Specifically we wanted to study whether response-format affected the results on the indirect measures. In experiment 1, at test participants had 4 alternative response buttons: “Old & Agree”, “Old & Disagree”, “New & Agree”, and “New & Disagree”. In experiment 2, participants were randomly assigned to one of two conditions: one group had the same response-types as in experiment 1; the other group had “Old”, “New”, “Agree”, and “Disagree” as alternatives to choose from at test.

Method

Participants

Participants were 61 university students in Sweden. Mean age was 23.82 with a standard deviation of 6.08. The sample consisted of 59% female and 41% male individuals.

Stimuli

Concept stimuli consisted of 60 IAPS images: 6 negative actions (e.g. an individual smoking crack), 6 neutral actions, 6 negative animals (e.g. a dog with exposed teeth), 6 neutral animals, 6 negative environmental scenes (e.g. a pile of garbage), 6 neutral environmental scenes, 6 negative humans (e.g. an angry individual), 6 neutral humans, 6 negative objects (e.g. an electric chair) and 6 neutral objects. All images are listed in the appendix. To establish that negative images were more negative than neutral images we calculated a set of repeated measures ANOVAs with average valence ratings from the IAPS manual as the dependent variable, and with images as cases. As shown in table 5, in each category, negative images were more negative than neutral images (averages in the manual are based on valence ratings on a 9-point scale where 1 = negative and 9 = positive).

Table 5.

Valence ratings for the image stimuli

Image type	*M* - Negative	*M* - Neutral	*F*	*p*	η^2
Actions	3.20	5.44	105.70	< 0.05	.96
Animals	3.79	6.84	164.20	< 0.05	.97
Environment	3.26	6.08	62.14	< 0.05	.93
Humans	3.56	5.03	14.18	< 0.05	.75
Objects	3.51	5.12	142.94	< 0.05	.97

Attribute stimuli consisted of 60 words: 6 negative words that could be true of actions (e.g. “despicable”), 6 neutral words that could be true of actions, 6 negative words that could be true of animals (e.g. “brutal”), 6 neutral words that could be true of animals, 6 negative words that could be true of environmental scenes (e.g. “dirty”), 6 neutral words that could be true of environmental scenes, 6 negative words that could be true of humans (e.g. “hostile”), 6 neutral words that could be true of humans, 6 negative words that could be true of objects (e.g. “dangerous”) and 6 neutral words that could be true of objects. All words are listed in the appendix. Negative and neutral words were matched on frequency of occurrence in the Swedish language. Negative and neutral words were not matched on word length however.

To establish that negative words are more negative than neutral words we collected ratings from 6 independent judges (5 – point scale, 1 = very negative, 5 = very positive). Repeated measures ANOVAs with the averages over judges as the dependent variable and words as cases showed that words had the expected valences (see table 6).

Table 6.

Valence ratings for the word stimuli

Presented with	*M* - negative	*M* - neutral	*F*	*p*	η^2
Actions	1.64	3.00	23.86	< 0.05	0.83
Animals	1.50	3.48	231.17	< 0.05	0.98
Environment	1.83	3.14	64.91	< 0.05	0.93
Humans	1.88	4.01	78.96	< 0.05	0.94
Objects	1.45	3.48	42.15	< 0.05	0.89

Type of word was nested within levels of type of image (e.g. negative words that could be true of humans, such as “angry”, were not presented together with images of objects, etc).

Procedure

Participants were tested using the standard IDAM as described in the introduction, with the exception that one group of participants had “Old”, “New”, “Agree” and “Disagree” as response alternatives at test. For actions, animals, environment, humans and objects, negative and neutral were concept categories with images as instances; for attribute categories, negative and neutral were concept categories with words as instances.

At study participants performed the relational encoding task used in experiment 1 (i.e. rate if an image + word combination on a given trial was similar to the image + word combination that had been shown on the previous trial). Again, ratings were made by clicking with the computer mouse on an alternative on a 4-point scale that ranged from “very dissimilar” to “very similar”, there being no time limit. The rationale for this task was to focus attention on the faces and words. The study phase consisted of 60 trials in total.

Before going through the test phase participants completed the 7-item vocabulary test described in experiment 1. The test phase then consisted of the standard IDAM for one group of participants (in the combined responses condition) with “Old & Agree”, “Old & Disagree”, “New & Agree” and “New & Disagree” as response alternatives. Response alternatives appeared in random positions on each trial. The other group (in the separate responses condition) had “Old”, “New”, “Agree” and “Disagree” as response alternatives. Response buttons for “Old” and “New” were activated first, “Agree” and “Disagree” – buttons were only enabled when participants had made an old/new response. In this condition buttons did not appear in random positions. Participants were allotted randomly to response conditions.

Results

Descriptive statistics for the response frequencies are shown in table 7 (congruent = negative image + negative word and neutral image + neutral word, incongruent = negative image + neutral word, neutral image + negative word). For the separate responses condition, a certain response category was scored when participants made two specific responses on a given trial (e.g. HA was scored when a participant clicked the “Old” – button on and clicked the “Agree” – button on an old combination).

Table 7.

Response frequencies

	M		*SD*	
	Congruent	Incongruent	Congruent	Incongruent
HA	0.50	0.22	0.15	0.12
FA	0.14	0.04	0.11	0.06
CA	0.52	0.16	0.16	0.10
MA	0.22	0.06	0.13	0.06
HD	0.21	0.51	0.09	0.14
FD	0.06	0.17	0.12	0.17
CD	0.28	0.63	0.11	0.20
MD	0.07	0.20	0.06	0.11

Next we calculated IDAM variables for congruent and incongruent combinations as described in the introduction: DM, IIM_L, EIM_L, IIM_{MN}, and EIM_{MN}. Like in experiment 1 we replaced outliers (values with a two-tailed probability lower than 0.01 in a normal distribution) with the values (raw scores) that correspond to 0.01/2 and 1 – 0.01/2. The variable that had the highest number of outliers had 8.20% outliers.

We started by evaluating the linear and multinomial algorithm regarding their ability to classify congruent and incongruent combinations, as well as looking at effects of response types. So we conducted two simpler ANOVAs without type of type of image as a factor, i.e. collapsing across actions, animals, environmental scenes, objects and humans. The first ANOVA, with the linear indirect measures, had congruence (2: congruent, incongruent) and measure (IIM_L, IIM_{MN}) as within subjects factors; group (2: combined responses, separate responses) was a between subjects factor. None of the effects involving group where significant, as shown in table 8. Congruence x measure was significant, however, indicating that this factor had opposing effects depending on whether the implicit or the explicit indirect measure was used. Follow up analyses indicated that congruent combinations produce higher values on IIM_L than incongruent combinations, and that incongruent combinations produced higher values on EIM_L than congruent combinations; see table 8 and figure 2. The linear implicit indirect attitude measure (IIM_L) could thus discriminate between congruent and incongruent combinations of images and words, with an effect size of 11% (see table 8).

When we used measures from the multinomial algorithm in a congruence x measure x group ANOVA, we found a similar pattern: no effects involving group, a significant interaction between congruence and measure, and a significant advantage for congruent combinations on IIM_{MN} but a significant advantage for incongruent combinations on EIM_{MN}; see table 8 and figure 1. Like IIM_L, IIM_{MN} could thus classify congruent and incongruent combinations. Note, however, that the effect of congruence on the multinomial indirect measure (IIM_{MN}) produced a higher effect size (20%).

Table 8.

Results on the indirect measures

Algorithm	Effect	*F*	*p*	η^2
Linear	Congruence x Measure x Group	0.20	0.66	0.00
	Congruence x Group	2.67	0.11	0.04
	Congruence x Measure	15.63	< 0.05	0.21
	Group	0.01	0.92	0.00
	Congruence	5.41	< 0.05	0.08
	Measure	14.98	< 0.05	0.20
	Congruence | IIM_L	7.39	< 0.05	0.11
	Congruence | EIM_L	19.37	< 0.05	0.24
Multinomial	Congruence x Measure x Group	0.01	0.94	0.00
	Congruence x Group	1.65	0.20	0.03
	Congruence x Measure	15.58	< 0.05	0.21
	Group	0.01	0.95	0.00
	Congruence	0.22	0.64	0.00
	Measure	1.73	0.19	0.03
	Congruence | IIM_{MN}	14.92	< 0.05	0.20
	Congruence | EIM_{MN}	10.27	< 0.05	0.15

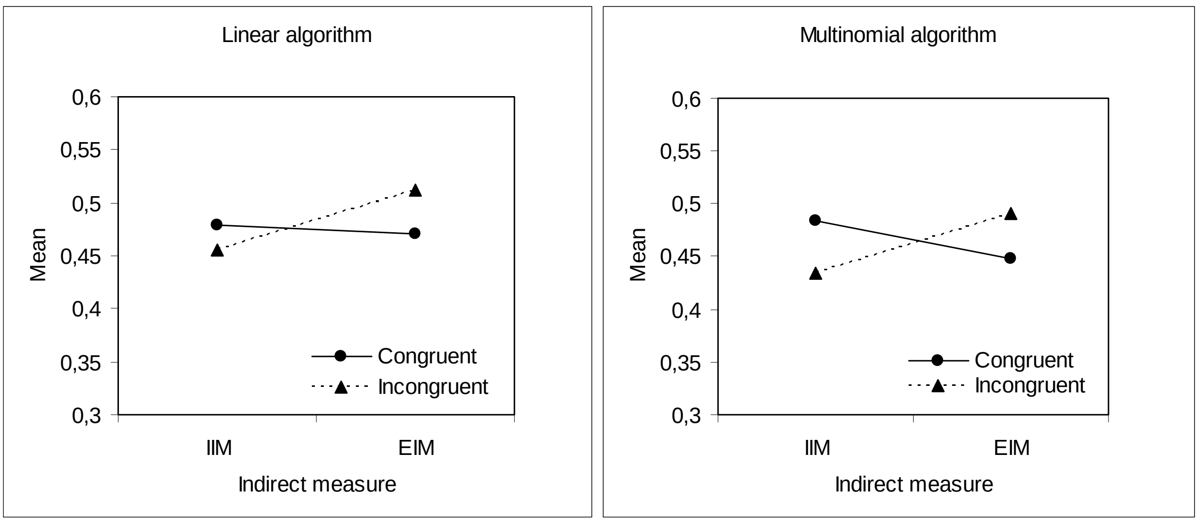

Figure 2. Averages for the experimental conditions on the indirect measures.

Congruence produced a large effect on the direct measure DM, F (1, 60) = 1131.39, $p < 0.05$, η^2 = 0.95, with a higher mean for congruent than for incongruent combinations (*M* = 0.59 and *SD* = 0.04, *M* = 0.37 and *SD* = 0.05, respectively).

In the following analysis we examined if IDAM could discriminate between congruent and incongruent combinations of different image types. Because IIM_{MN} produced the largest effect size we only used this measure as the dependent variable. Calculating IIM_{MN} resulted in division by zero in 2.13% of the data; these observations were imputed as missing values. Outliers were defined and treated in the usual way (the highest number of outliers amounted to 8.20 %). We then calculated a repeated measures ANOVA with congruence (2: congruent, incongruent) and image type (5: actions, animals, environment, humans, objects) as within subjects factors (Huynh – Feldt corrected p-values). This analysis showed a non-significant interaction between image type and congruence, F (4, 192) = 1.23, p = 0.30, η^2 = 0.03, and non-significant main effect of image type, F (4.00, 191.90) = 1.98, p = 0.10, η^2 = 0.04. The main effect of congruence was significant, however, F (1, 48) = 10.72, $p < 0.05$, η^2 = 0.18. As shown in figure 3, *all* congruent combinations produced higher values on the multinomial implicit indirect measure (IIM_{MN}) than *all* incongruent combinations.

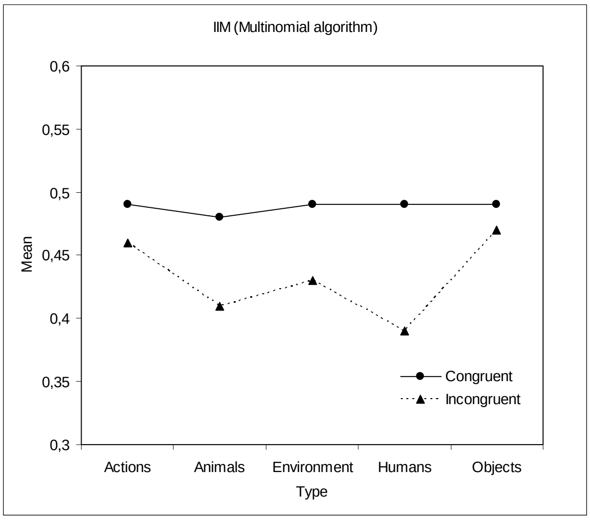

Figure 3. Averages for the experimental conditions and image types on the implicit indirect measure IIM.

Discussion

Like in experiment 1, the implicit indirect measures from IDAM could discriminate between congruent and incongruent combinations, with the multinomial algorithm (IIM_{MN}) producing the largest effect size. Experiment 2 involved a wide variety of stimuli: everything from crack addicts to sharks. In spite of this, the implicit indirect measure from the multinomial algorithm showed higher values on congruent than incongruent combinations. This evidence further strengthens the measures construct validity as an indicator of implicit attitudes.

Experiment 2 also revealed that incongruent combinations had higher values on the explicit indirect measures (EIM_L and EIM_{MN}). Again, this shows that implicit and explicit attitudes are dissociated, supporting the usefulness of the implicit attitude construct. Like in experiment 1, we can explain this phenomenon by assuming that fluency is greater for congruent than for incongruent combinations and by assuming that fluency is only attributed to memory when participants make “disagree” responses (when they make “agree” responses on congruent combinations there is no reason to attribute fluency to “old”). Response format, combined or separate, did not affect the pattern of results. It is reassuring that the IDAM is insensitive to this kind of superficial manipulation.

Experiment 3

The purpose of experiment 3 was to examine effects of different study tasks. Experiments 1 and 2 both used relational encoding instructions: Participants had to judge whether a combination on a given trial was similar to the combination on the previous trial, by making a rating on a four – point scale that had “very similar” and “very dissimilar” as end points. The purpose of this task was to focus attention on the stimuli in experiments 1 and 2. But the indirect measures of IDAM are not based on the ability to discriminate between old and new combinations; instead they rely on recognition biases. Whether a participant pays careful attention or not to concept instances + attribute instances should only influence recognition sensitivity and not recognition bias. In experiment 3 we examined if we could do away with the relational encoding task used in experiments 1 and 2; if so, IDAM – testing would be more time efficient.

Method

Participants

Sixty two high school students and university students from Sweden participated in the experiment; mean age was 21.11 (*SD* = 6.84). The sample consisted of 66% female and 44% male individuals.

Stimuli

Concept instances were exactly the same IAPS images used in experiment 2: 6 negative actions, 6 neutral actions, 6 negative animals, 6 neutral animals, 6 negative environmental scenes, 6 neutral environmental scenes, 6 negative humans, 6 neutral humans, 6 negative objects and 6 neutral objects. Attribute instances were the words used in experiment 2: 6 negative action words, 6 neutral action words, 6 negative animal words, 6 neutral animal words, 6 negative environment words, 6 neutral environment words, 6 negative human words, 6 neutral human words, 6 negative objects words, and 6 neutral object words. Like in experiment 2, type of word was nested within levels of type of image. Images and words are listed in the appendix.

Procedure

Participants went through the standard IDAM as described in the introduction. At study participants were randomly assigned to one of two conditions: Relational or simple. In the relational condition, participants performed exactly the same task as in experiments 1 and 2 (i.e. rate if a combination was similar to the combination on the previous trial). In the simple condition participants were simply asked to watch the images and the words carefully. The study phase consisted of 60 trials. After the study phase participants conducted the test phase which consisted of 120 trials (here we thus left out the distraction task that was used in experiments 1 and 2).

Results

Response frequencies were calculated for congruent (negative image + negative word, neutral image + neutral word) and incongruent combinations (negative image + neutral word, neutral image + negative word). The associated descriptive statistics are shown in table 9.

Table 9.

Response frequencies

	M		*SD*	
	Congruent	Incongruent	Congruent	Incongruent
HA	0.45	0.21	0.16	0.13
FA	0.19	0.08	0.14	0.08
CA	0.45	0.16	0.16	0.10
MA	0.23	0.09	0.11	0.07
HD	0.19	0.45	0.10	0.15
FD	0.08	0.19	0.11	0.18
CD	0.28	0.56	0.12	0.21
MD	0.13	0.24	0.08	0.14

Then we calculated IDAM variables as described in the introduction, i.e. DM, IIM_L, IIM_{MN}, EIM_L, EIM_{MN} (outliers, max 4.84%, were treated as always). Using the indirect measures from the linear algorithm as the dependent variables, we computed an ANOVA with Congruence (2: congruent, incongruent), Measure (2: IIM_L, EIM_L) and Group (2: relational, simple). The results are shown in table 10. None of the effects involving group were significant. Apparently, doing the simple or the relational study task did not have any impact on the implicit measures. The interaction between congruence and measure was also not significant, which is a bit surprising since such an interaction was evident in experiments 1 and 2. Note also that the linear indirect implicit measure (IIM_L) could not significantly differentiate between congruent and incongruent combinations. EIM_L was significant, though, with higher values for incongruent than congruent combinations (see table 10 and figure 4). In sum, the linear implicit indirect measure was unsuccessful at discriminating between congruent and incongruent combinations.

We then turned to the multinomial indirect measures, calculating an ANOVA with Congruence, Measure (IIM_{MN}, EIM_{MN}) and Group as factors. As before, none of the effects of group were significant (see table 10). The interaction between congruence and measure proved to be reliable, however. Follow up analyses indicated that congruent combinations had higher values on IIM_{MN} than incongruence combinations; congruence could explain about 8% of the variance in IIM_{MN}. Congruent combinations were not significantly different from incongruent combinations on EIM_{MN} (see table 10 and figure 4). So the multinomial implicit indirect measure (IIM_{MN}) succeeded in differentiating congruent from incongruent combinations, in the expected direction. Note that the magnitude of this effect was rather low, however. As with the linear algorithm, group did not have any influence at all on indirect implicit and explicit attitude measures.

Table 10.

Results for the indirect measures

Algorithm	Effect	*F*	*p*	η^2
Linear	Congruence x Measure x Group	0.06	0.81	0.00
	Congruence x Group	0.17	0.68	0.00
	Congruence x Measure	2.70	0.11	0.04
	Group	0.10	0.76	0.00
	Congruence	0.13	0.29	0.02
	Measure	23.52	< 0.05	0.28
	Congruence | IIM_L	0.65	0.42	0.01
	Congruence | EIM_L	4.03	< 0.05	0.06
Multinomial	Congruence x Measure x Group	0.08	0.78	0.00
	Congruence x Group	0.00	0.97	0.00
	Congruence x Measure	5.56	< 0.05	0.10
	Group	0.24	0.63	0.00
	Congruence	0.50	0.48	0.01
	Measure	13.48	< 0.05	0.18
	Congruence | IIM_{MN}	5.17	< 0.05	0.08
	Congruence | EIM_{MN}	2.80	0.10	0.04

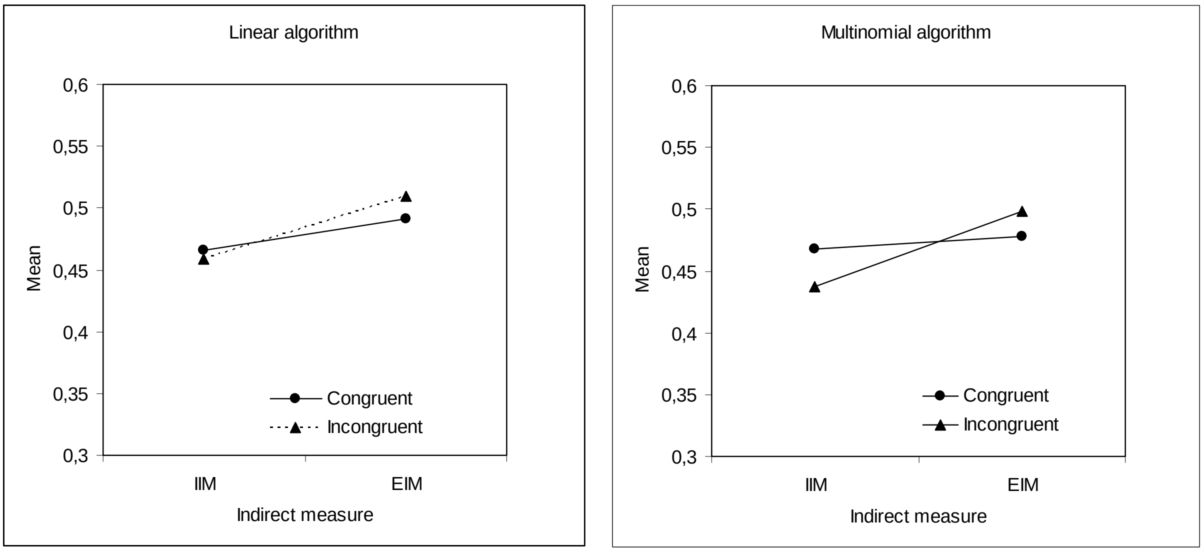

Figure 4. Means for conditions on the indirect measures.

On the direct measure (DM) there was a large main effect of congruence, F (1, 61) = 794.89, $p < 0.05$, $\eta^2 = 0.93$, indicating that congruent combinations had higher values than incongruent combinations (M = 0.58 and SD = 0.05, M = 0.39 and SD = 0.06, respectively). We are not surprised by this result.

Discussion

The most important finding from experiment 3 was that the linear implicit indirect measure (IIM_L) failed to discriminate between congruent and incongruent combinations, and that the multinomial implicit indirect measure (IIM_{MN}) succeeded at discriminating between congruent and incongruent combinations. Bu this effect size only amounted to 8%, which is considerably lower than the corresponding effect sizes in experiment 1 (31%) and experiment 2 (20%). An explanation of this discrepancy is that the sample in this experiment consisted of a number of high school students; they might not have taken the task as seriously as the university students that participated in experiments 1 and 2.

Type of study task (relational or simple) did not affect the outcome on the multinomial implicit indirect measure (IIM_{MN}). This is good, since the relational study task in experiments 1 and 2 took quite a lot of time. If the IDAM is going to be used in applied situations, testing time is an important factor.

Experiment 4

Experiment 3, in which concept instances consisted of affective images, revealed that we could dispense of the relational study task. The purpose of experiment 4 was to examine if the same would be true when conducting the IDAM to measure attitudes towards physically attractive and physically unattractive individuals, like in experiment 1. Because experiment 3 did not reveal any significant effects of study task (relational or simple) we did not manipulate this variable here. Instead, all participants went through the simple study task, with the instruction to watch the images and words carefully.

Method

Participants

Participants were 30 university students in Sweden; mean age was 23.67 (*SD* = 4.61). We tested 63% female and 37% male subjects.

Stimuli

Stimuli were the same as in experiment 1: Instances for the concept category *attractive* were 20 attractive faces, instances for the concept category *unattractive* were 20 unattractive faces, instances for the attribute category *positive* were 20 positive words, and instances for the attribute category *negative* were 20 negative words. Images and words are listed in the appendix.

Procedure

The procedure consisted of the standard IDAM. The study phase consisted of 40 trials and participants were instructed to simply watch each face and word carefully. The test phase consisted of 80 trials.

Results

Response frequencies were calculated for congruent combinations (attractive + positive and unattractive + negative) and incongruent combinations (attractive + negative and unattractive + positive); descriptive measures for each condition are shown in table 11.

Table 11.

Response frequencies

	M		*SD*	
	Congruent	Incongruent	Congruent	Incongruent
HA	0.36	0.22	0.17	0.17
FA	0.18	0.08	0.12	0.08
CA	0.37	0.19	0.19	0.14
MA	0.26	0.12	0.13	0.09
HD	0.18	0.34	0.08	0.16
FD	0.14	0.24	0.16	0.21
CD	0.32	0.49	0.18	0.21
MD	0.20	0.32	0.15	0.16

We then computed the standard IDAM variables for congruent and incongruent combinations: DM, IIM_L, EIM_L, IIM_{MN}, and EIM_{MN}. The variable that had the highest number of outliers had 3.33% outliers; these were treated according to the usual method.

ANOVAs with congruence (2: congruent and incongruent) and measure (2: IIM and EIM) were computed for the linear and multinomial indirect measures. Table 12 gives a rather disconcerting impression: None of the implicit indirect measures could discriminate between congruent and incongruent combinations and the effect size for IIM_{MN} only amounted to 7%. The effect of congruence on IIM_{MN} is, although, in the predicted direction.

Table 12.

Results on the indirect measures

Algorithm	Effect	*F*	*p*	η^2
Linear	Congruence x Measure	0.67	0.42	0.02
	Congruence	0.50	0.49	0.02
	Measure	6.15	< 0.05	0.18
	Congruence | IIM_L	0.13	0.73	0.00
	Congruence | EIM_L	1.04	0.32	0.03
Multinomial	Congruence x Measure	2.73	0.11	0.09
	Congruence	0.08	0.78	0.00
	Measure	2.74	0.11	0.09
	Congruence | IIM_{MN}	2.00	0.17	0.07
	Congruence | EIM_{MN}	1.77	0.19	0.06

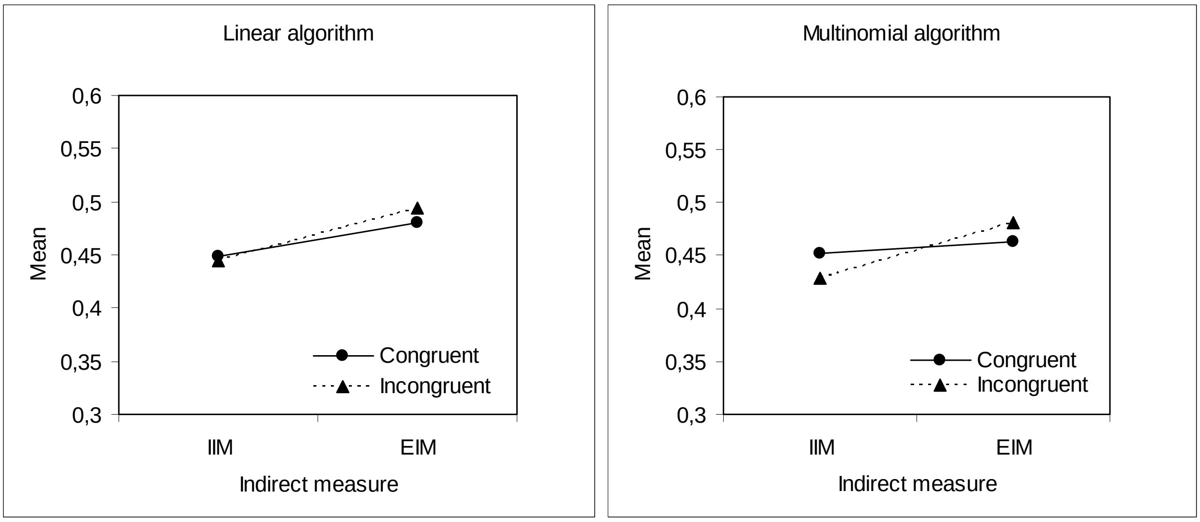

Figure 5. Results for conditions on the indirect measures.

On the direct measure (DM) there was a large main effect of congruence, $F(1, 29) = 179.24$, $p < 0.05$, $\eta^2 = 0.86$. Congruent combinations had higher values than incongruent combinations ($M = 0.55$ and $SD = 0.04$, $M = 0.40$ and $SD = 0.06$, respectively).

Discussion

The results of experiment 4 were puzzling: experiment 1 that used exactly the same stimuli found a rather large effect (31%) of congruence on the multinomial implicit indirect measure (IIM_{MN}); this effect was non-significant in experiment 4 even though it was in the expected direction. The only difference between experiment 1 and experiment 4, was that experiment 1 used relational study instructions (i.e. rate if this combination is similar to the previous combination) while experiment 4 used the simple study instruction (i.e. pay attention to images and words). But when we examined this factor in experiment 3, type of study instruction did not influence the implicit indirect measures at all. So this cannot be a viable explanation. At this point we can only point to individual differences in the two samples as an account.

Experiment 5

In experiment 5 we wanted to study if repeating specific instances of concept categories across trials or repeating specific instances of attribute categories across trials would have any impact on the measurement of implicit attitudes in IDAM. This may be necessary in certain circumstances. Assume that we want to measure implicit attitudes towards a certain type of car, car X. In this case it might be difficult to find several instances of the concept category car X; i.e. we might have to rely on say 2 different images of the same car. But for reliability reasons, we would have to present several trials involving car X; e.g. image 1 of car X + the word “nice”, image 1 of car X + the word “beautiful”, image 1 of car X + the word “attractive”, image 2 of car X + the word “ugly”, image 2 of car X + the word “unattractive”, image 2 of car X + the word “hideous”, etc. Each combination would be unique, but an instance of the concept category (an image of the car) is repeated several times. In the same way, we might want to use IDAM in a study in which instances of attribute categories have to be repeated; e.g. we might want to use general positive or negative words, such as “positive”, “good”, “negative” and “bad” in conjunction with two concept categories (e.g. female and male). But for reliability reasons attribute words would have to be repeated; so the word “positive” would have to appear alongside several different images of different female individuals, for example. In other cases it might be interesting to study if individuals associate a certain concept with a very narrow attribute category, say the concept category car X with the attribute category fast; there is a limited number of instances that represent the category fast.

Combinations of concept instance + attribute instance that involve repeated concept instances or repeated attribute instances are less distinctive than combinations that do not involve repeated instances. It is well known that distinctiveness affects recognition sensitivity; i.e. unique items are easier to remember than non-unique items (Tulving & Thomson, 1973). Distinctiveness might affect recognition bias as well; the indirect measures of IDAM heavily rely on recognition bias. When specific words or specific images are repeated, even if a combination of image and word is unique, the recognition memory task becomes and is experienced as more difficult. If participants experience that they have difficulties in identifying old items correctly, they might be less prone to use fluency as a sign of memory and instead rely more on right out guessing. This would obviously pose a problem when measuring implicit attitudes.

In experiment 5 we manipulated distinctiveness between subjects: one group of subjects received an IDAM with unique words and unique images in each trial at study and test. Another group of subjects received an IDAM with repeated words or repeated images across trials at study and test (even though image + word combinations were unique).

Method

Participants

Participants were 61 university students from Sweden. Mean age was 24.15 (*SD* = 6.65). There were 67% female and 33% male individuals in the sample.

Stimuli

Stimuli were the same as in experiments 1 and 4: 20 unattractive faces, 20 attractive faces, 20 negative words and 20 positive words (see the appendix).

Procedure

Participants completed the standard IDAM as described in the introduction. They were randomly assigned to one of three conditions. (1) In the distinctive condition (N = 31), each one of the 20 attractive and 20 unattractive faces was paired with a unique negative or positive word; i.e. there was one word per face. (2) In non-distinctive condition 1 (N = 15) only 2 attractive and 2 unattractive faces were paired with the 20 negative and 20 positive words; these faces were randomly selected for each participant from the set of 20 attractive and 20 unattractive faces. (3) In non-distinctive condition 2 (N = 15) only 2 negative and 2 positive words were paired with the 20 attractive and 20 unattractive faces; these words were randomly selected for each participant from the set of 20 negative and 20 positive words. The conditions are shown schematically in table 13.

Table 13.

Schematic description of between - subjects conditions

Condition	
Distinctive	Concept A: Instance A1 + Attribute L: Instance L1
	Concept A: Instance A2 + Attribute L: Instance L2
	Concept B: Instance B1 + Attribute L: Instance L3
	…
Non-distinctive 1	Concept A: Instance A1 + Attribute L: Instance L1
	Concept A: Instance A1 + Attribute L: Instance L2
	Concept B: Instance B1 + Attribute L: Instance L3
	…
Non-distinctive 2	Concept A: Instance A1 + Attribute L: Instance L1
	Concept A: Instance A2 + Attribute L: Instance L1
	Concept B: Instance B1 + Attribute L: Instance L2
	…

The study phase consisted of 40 trials and participants completed the task of judging whether a combination was similar to the combination on the previous trial (experiment 3 showed that this task did not affect implicit attitudes but the results from experiment 3 were not available when we planned experiment 5). Before completing the test phase participants went through the 7-item word comprehension test. The test phase consisted of 80 trials.

Results

Averages and standard deviations of response frequencies for congruent and incongruent combinations are shown in table 14. Congruent combinations consisted of attractive + positive and unattractive + negative. Incongruent combinations consisted of attractive + negative and unattractive + positive.

Table 14.

Response frequencies

	M		*SD*	
	Congruent	Incongruent	Congruent	Incongruent
HA	0.34	0.21	0.15	0.15
FA	0.21	0.11	0.17	0.12
CA	0.30	0.17	0.15	0.11
MA	0.23	0.14	0.15	0.10
HD	0.21	0.34	0.12	0.16
FD	0.16	0.22	0.12	0.19
CD	0.33	0.49	0.18	0.22
MD	0.21	0.31	0.12	0.15

Next we computed IDAM variables for congruent and incongruent combinations: DM, IIM_L, EIM_L, IIM_{MN}, EIM_{MN} and outliers were defined and treated as in experiments 1-4. The variable that had the highest number of outliers had 4.92% outliers.

We then conducted a repeated measures ANOVA with Group (2: distinctive, non-distinctive 1 + 2), Congruence (2: congruent, incongruent) and Measure (2: IIM_{MN}, EIM_M). We put non-distinctive 1 + 2 together because we were mainly interested in the effects of using distinctive versus non-distinctive combinations, and not in the effects of using non-distinctive-faces versus non-distinctive words. We also only calculated measures from the multinomial algorithm for space reasons. This analysis revealed a significant interaction between group, congruence and measure and a significant main effect of group as shown in table 15. The main effect of group indicated that participants in the non-distinctive condition had higher values on the indirect measures as shown in figure 6.

Separate follow up ANOVAs for each group, revealed that the effect of congruence x measure was significant in the distinctive group but not significant in the non-distinctive group, and that the main effect of congruence was non-significant in both groups, see table 15 and figure 6. In the distinctive group, we found a significant effect of congruence on the implicit indirect measure (IIM_{MN}), with higher values on congruent than on incongruent combinations, see table 15 and figure 6; congruence could explain 0.21% of the variance in IIM_{MN} here. The effect of congruence did not reach significance for the explicit indirect measure (EIM_{MN}) in the distinctive group.

Table 15.

Results on the indirect measures (calculated with the multinomial algorithm)

Group	Effect	*F*	*p*	η^2
Both	Congruence x Measure x Group	8.35	< 0.05	0.12
	Congruence x Group	0.91	0.34	0.02
	Group	12.46	< 0.05	0.17
Non-distinctive (1 + 2)	Congruence x Measure	0.60	0.45	0.02
	Congruence	4.11	0.05	0.12
	Measure	6.36	< 0.05	0.18
Distinctive	Congruence x Measure	10.94	< 0.05	0.27
	Congruence	0.76	0.38	0.03
	Measure	11.98	< 0.05	0.29
	Congruence | IIM_{MN}	7.90	< 0.05	0.21
	Congruence | EIM_{MN}	4.11	0.05	0.12

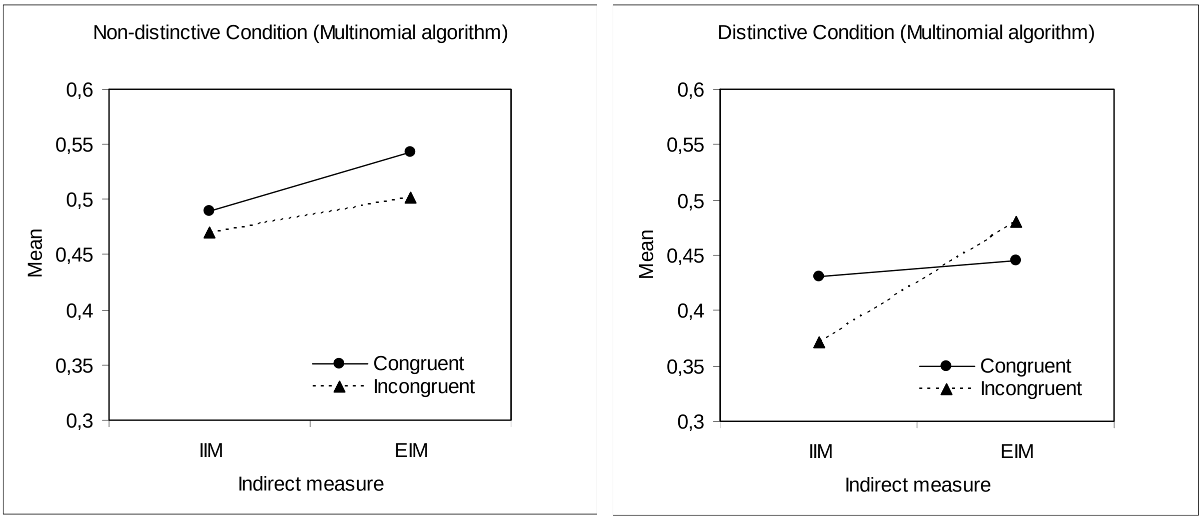

Figure 6. Results on the indirect measures for the two groups.

Finally we computed an ANOVA with congruence and group as factors and the direct measure (DM) as the dependent variable. This analysis did not show an effect of congruence x group, F (1, 59) = 0.49, p = 0.49, eta2 = 0.01. The main effects of congruence and group were significant, however, F (1, 59) = 66.91, p < 0.05, eta2 = 0.53, and F (1, 59) = 6.60, p < 0.05, eta2 = 0.10, respectively. Congruent combinations had higher values on DM than incongruent combinations (M = 0.52 and SD = 0.08, M =

0.41, *SD* = 0.09, respectively). Participants in the non-distinctive group had higher values on DM than participants in the distinctive group (*M* = 0.49 and *SD* = 0.10, *M* = 0.45, *SD* = 0.10, respectively).

Discussion

Experiment 5 demonstrated that the implicit indirect measure only discriminated between congruent and incongruent combinations when stimulus items were distinct; i.e. when each face vas paired with a different word. When using the IDAM in a certain application this factor is thus important to consider. Another finding was that participants in the non-distinctive group generally had higher values on the indirect measures (IIM_{MN} and EIM_{MN}). Items in the non-distinctive condition are more similar to each other and they should be more easily confounded, probably prompting more guessing responses.

General discussion

Summary of findings

The purpose of the present research was to develop and evaluate a new indirect measure of attitudes that can yield valid measurements of implicit attitudes as introspectively unidentified associations. We examined two algorithms to calculate implicit attitudes, a number of procedural variations in the test, and whether the IDAM could measure implicit attitudes for a number of different stimulus types.

The multinomial algorithm for computing implicit attitude (IIM_{MN}) performed better than the linear algorithm for computing implicit attitude (IIM_{L}). In each experiment that examined this factor, congruence produced higher effect sizes on IIM_{MN} than on IIM_{L}. Regarding IIM_{MN}, this measure could significantly discriminate between congruent and incongruent combinations in 4 of 5 experiments; experiment 4 failed to reveal a significant effect of congruence on IIM_{MN} however. In order to arrive at a general conclusion about the validity of IDAM we calculated an ANOVA with Congruence (2: congruent, incongruent) as a within subjects factor and Experiment (5) as between subjects factor. The implicit indirect measure from the multinomial algorithm (IIM_{MN}) was the dependent variable (in this analysis we excluded participants that received non-distinctive combinations in experiment 5). This analysis showed that the interaction between congruence and experiment was non-significant, F (4, 211) = 1.00, $p = 0.41$, $\eta^2 = 0.02$. Means for conditions and experiments are presented in figure 7; the difference between congruent and incongruent conditions was thus not moderated by experiment. As expected there was a main effect of congruence, F (1, 211) = 36.90, $p < 0.05$, $\eta^2 = 0.15$. Across experiments, congruent combinations produced higher values on IIM_{MN} than incongruent combinations. Finally, there was a significant main effect of experiment, F (4, 211) = 2.66, $p < 0.05$, eta2 = 0.05, indicating that the magnitude of IIM_{MN} varied quadratically across experiments (see figure 7). The absence of an interaction between congruence and experiment and the main effect of congruence are important. Provided that distinctive concept instance + attribute instance combinations are used, the IDAM seems to be a valid indirect indicator of implicit attitude.

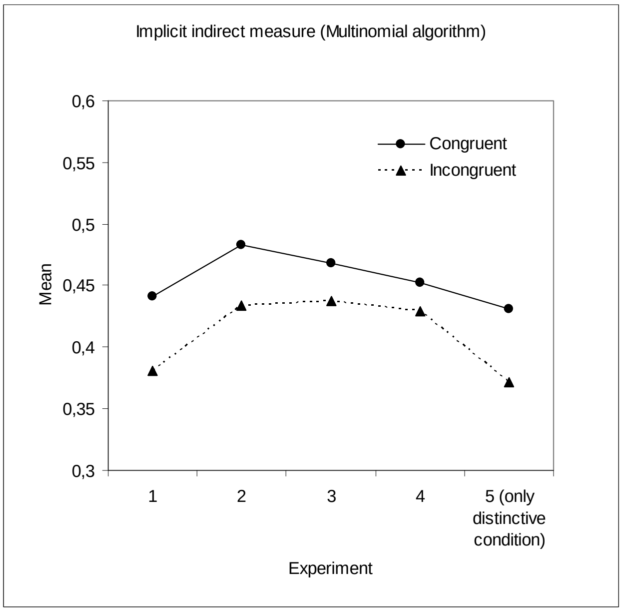

Figure 7. Averages for congruent and incongruent combinations for experiments 1-5. Note that participants that received non-distinctive combinations in experiment 5 were excluded.

Even so, effect sizes for the multinomial implicit indirect measure varied considerably between experiments as shown in the summary in table 16: From 6% explained variance to 31% explained variance (excluding participants that received non-distinctive items in experiment 5). There is no obvious explanation for this fact. Both attractive and unattractive faces and IAPS images sometimes produced larger effect sizes (e.g. experiments 1 and 2) and sometimes produced small effect sizes (e.g. experiments 3 and 4). The same was true for type of study task (relational or simple) and distraction task (present or absent); see table 16. Note also that female/male ratio and mean age do not appear to correlate systematically with effect sizes. At this point we can only refer to individual differences between participants in the different samples as a tentative explanation. We know, however, that non-distinctive items nullify the effect of congruence on the, multinomial implicit indirect measure (see experiment 5).

Table 16.

Overview of effect sizes for experiments 1-5

Exp	Distinctive combinations	Stimuli	Study task	Distraction task	Female/male ratio	Mean age	IIM_{MN} effect size
1	Yes	Faces	Relational	Yes	69/31	25.38	0.31
5	Yes	Faces	Relational	Yes	68/32	23.71	0.21
2	Yes	Images	Relational	Yes	59/41	23.82	0.20
3	Yes	Images	Simple	No	65/35	21.42	0.10
4	Yes	Faces	Simple	No	63/37	23.67	0.07
3	Yes	Images	Relational	No	69/31	20.81	0.06
5	No	Faces	Relational	Yes	67/33	24.60	0.03

Each experiment revealed an interaction between congruence and measure: Congruent combinations had an advantage over incongruent combinations but only when the indirect measure was conditional on the direct measure showing evidence that there was no attitude. On the explicit indirect measure the effect was reversed and sometimes absent. This observation is important, because it indicates the importance of distinguishing implicit attitudes from explicit attitudes.

Interpretations of the findings

The multinomial indirect measure of implicit attitude (IIM_{MN}) was better than the linear indirect measure of implicit attitude. An explanation of this fact is that the linear measure might be confounded by recollection, and not perform as a pure measure of responding "old" and not "new" (bias). The linear implicit indirect measure, IIM_L, was defined as HD + FD - CD - MD. If we assume that HD = recollection + bias, that FD = bias, that CD = recollection – bias, and that MD = - recollection - bias, we get IIM_L = 4 (bias) + recollection. I.e. a component of IIM_L is correlated with the ability to remember old and new combinations. In the multinomial measure, instead, measures of recognition accuracy, e.g. (- 1 - FD + HD) or (- 1 + CD - MD), always constitute the divisor. This means that each term in the equation of IIM_{MN} is held constant across different levels of memory performance.

Experiment 5 demonstrated an effect on the multinomial implicit indirect measure only when participants conducted the IDAM with distinctive items in which each face was paired with a unique word. An explanation of this phenomenon is that participants experience the task as to difficult when items are non-distinctive; they might then fall back on simple guessing. The IDAM relies on fluency for congruent combinations. Fluency is not the same as guessing; in order to experience fluency an individual must feel a certain degree of confidence in his or her responses; i.e. fluency has to be genuinely attributed to memory. When guessing, a participant instead probably attributes his or her "old"-responses to chance or not knowing.

Each experiment found an interaction between congruence and type of indirect measure (implicit or explicit). A reasonable explanation for this phenomenon is that only congruent combinations produce fluency and that the fluency that is produced by congruent combinations is only attributed to "old" when it is not attributed to "agree". I.e. the feeling that a congruent item is easy to process does not have to be attributed to "old" when it can be attributed to "agree".

Suggestions for further research

The present experiments generally supported the IDAMs construct validity by showing that it could discriminate between congruent and incongruent combinations. Nevertheless it might be a bit premature to start using the technique in applied settings or in other basic research. For instance it would be reassuring to find that the technique can be used to predict some form of behaviour. In fact, we conducted such a study, but this IDAM employed non-distinctive combinations of concept instances and attribute instances (experiment 5 showed that the IDAM does not seem to work with non-distinctive items). In this experiment we measured attitudes towards six different colours of mp3 – players, with positive and negative words as attribute instances. A certain image of an mp3 player was repeated a number of times together with different words. After having performed the IDAM, participants had to choose a flyer with an image of a certain colour of mp3 - player. In a lottery they could then win an mp3 – player that had the colour that they chosen. The results indicated that implicit measures from the IDAM could *not* predict choice of flyer. In the light of the findings from experiment 5 it is difficult to draw any conclusions regarding if IDAMs implicit measures can predict behaviour or not. Had we conducted an IDAM with distinctive stimuli, we might have been able to make such a prediction.

Conclusions

Based on the findings from the present experiments we are rather confident in concluding that the IDAM is a valid measure of implicit attitude. The measure could discriminate between several different types of attitude objects. The jury is still out, however, regarding if the IDAM has predictive validity in relation to behaviour or not.

References

Eagly, A. H., & Chaiken, S. (1998). Attitude structure and function. In D. T. Gilbert, S. T. Fiske & G. Lindzey (Eds.), The handbook of social psychology, Vol. 2 (4th ed.). (pp. 269-322): McGraw-Hill, New York, NY, US.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. Psychological Review, 102 (1), 4–27.

Bornstein, R. F. (1992). Subliminal mere exposure effects. In R. F. Bomstein & T. S. Pittman (Eds.), Perception without awareness: Cognitive, clinical, and social perspectives (pp. 191–210). New York: Guilford Press.

Ajzen, I. (1988). Attitudes, personality, and behavior: Dorsey Press, Homewood, IL, US.

Allport, G. W. (1935). Attitudes. In C. Murchinson (Ed.), A handbook of social psychology (pp 798–844). Worchester, MA: Clark University Press.

Banaji, M. R., & Greenwald, A. G. (1994). Implicit stereotyping and unconscious prejudice. In M. P. Zanna & J. M. Olson (Eds.), The psychology of prejudice. The Ontario Symposium (Vol. 7, pp 55–76). Hillsdale, NJ: Erlbaum.

De Houwer, J., & Moors, A. (2007). How to Define and Examine the Implicitness of Implicit Measures. In Wittenbrink, Bernd; Schwarz, Norbert (2007). Implicit measures of attitudes. (pp. 179–194). New York, NY, US: Guilford Press. x, 294 pp. New York, NY: Guilford Press.

Dovidio, J. F., Evans, N., & Tyler, R. B. (1986). Racial stereotypes: The contents of their cognitive representations. Journal of Experimental Social Psychology, 22 (1), 22–37.

Bartlett, F. C. (1932). Remembering: A study in experimental and social psychology: Cambridge University Press, New York, NY, US.

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. Journal of Personality and Social Psychology, 62 (6), 893–912.

Blair I. 2001. Implicit stereotypes and prejudice. In Cognitive Social Psychology: The Princeton Symposium on the Legacy and Future of Social Cognition, ed. GB Moskowitz, pp. 359–74.

Mahwah, NJ: Erlbaum Brauer M, Wasel W, Niedenthal P. 2000. Implicit and explicit components of prejudice. Rev. Gen. Psychol. 4: 79–101

Dovidio JF, Kawakami K, Beach KR 2001. Implicit and explicit attitudes: examination of therelationship between measures of intergroup bias. In Blackwell Handbook of Social Psychology: Intergroup Processes, ed. R Brown and SL Gaertner, pp 175–197. Malden, MA:Blackwell

Eagly, A. H., & Chaiken, S. (1998). Attitude structure and function. In D. T. Gilbert & S. T. Fiske & G. Lindzey (Eds.), The handbook of social psychology, Vol. 2 (4th ed.). (pp. 269–322): McGraw-Hill, New York, NY, US.

Eagly, A. H., Chen, S., Chaiken, S., & Shaw-Barnes, K. (1999). The impact of attitudes on memory: An affair to remember. Psychological Bulletin, 125 (1), 64–89.

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

Fishbein, M., & Ajzen, I. (1974). Attitudes towards objects as predictors of single and multiple behavioral criteria. Psychological Review, 59–74.

Gaertner, S. L., & McLaughlin, J. P. (1983). Racial stereotypes: Associations and ascriptions of positive and negative characteristics. Social Psychology Quarterly, 46 (1), 23–30.

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

Greenwald, A. G. (1990. What cognitive representations underlie social attitudes? Bulletin of the psychonomic society, 28, 254–260.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. Journal of Personality and Social Psychology, 74 (6), 1464–1480.

Greenwald, A. G., Nosek, B. A., Banaji, M. R., & Klauer, K. C. (2005). Validity of the Salience Asymmetry Interpretation of the Implicit Association Test: Comment on Rothermund and Wentura (2004). Journal of Experimental Psychology: General, 134 (3), 420–425.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. Journal of Personality and Social Psychology, 50 (2), 229–238.

Hastie R. 1984. Causes and effects of causal attribution. J. Pers. Soc. Psychol. 46: 44–56

Hense, R. L., Penner, L. A., & Nelson, D. L. (1995). Implicit memory for age stereotypes. Social Cognition, 13 (4), 399–415.

Jacoby, L. L., & Kelley, C. M. (1992). A process-dissociation framework for investigating unconscious influences: Freudian slips, projective tests, subliminal perception, and signal detection theory. Current Directions in Psychological Science, 1 (6), 174–179

Kelley, C. M., & Rhodes, M. G. (2002). Making sense and nonsense of experience: Attributions in memory and judgment. In D. L. Medin (Ed.), The psychology of learning and motivation: Advanes in research and theory (Vol. 41,. San Diego, CA: Academic Press.

Maass A, Salvi L, Arcuri L, Semin GR. 1989. Language use in intergroup contexts: the linguistic intergroup bias. J. Pers. Soc. Psychol. 57: 981–93

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

Nosek, B. A. (2005). Moderators of the Relationship Between Implicit and Explicit Evaluation. Journal of Experimental Psychology: General, 134 (4), 565–584.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. American Psychologist, 17, 776–783.

Roediger, H. L., III, Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. Psychonomic Bulletin & Review, 8 (3), 385–407.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21 (4), 803–814.

Roediger, H. L. (1990). Implicit memory: Retention without remembering. American Psychologist, 45 (9), 1043–1056.

Schacter, D. L. (2001). The seven sins of memory: How the mind forgets and remembers: Houghton Mifflin Co, Boston, MA, US.

Sekaquaptewa D, Espinoza P, Thompson M, Vargas P, von Hippel W. 2002. Stereotype explanatory bias: Implicit stereotyping as a predictor of discrimination. J. Exp. Soc. Psychol. In press

Sinclair, L., & Kunda, Z. (1999). Reactions to a Black professional: Motivated inhibition and activation of conflicting stereotypes. Journal of Personality and Social Psychology, 77 (5), 885–904.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. Journal of Experimental Psychology: General, 117 (1), 34–50.

Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. Psychological Bulletin, 111 (1), 42–61.

Tedeschi, J. T., Schlenker, B. R., & Bonoma, T. V. (1971). Cognitive dissonance: Private ratiocination or public spectacle? American Psychologist, 26, 685–695.

Warrington, E. K., & Weiskrantz, L. (1968). A study of learning and retention in amnesic patients. Neuropsychologia, 6 (3), 283–291.

Wegner, D. M., & Bargh, J. A. (1998). Control and automaticity in social life. In D. T. Gilbert & S. T. Fiske & G. Lindzey (Eds.), The handbook of social psychology, Vol. 1 (4th ed.). (pp. 446–496): McGraw-Hill, New York, NY, US.

Von Hippel, W., Sekaquaptewa, D., & Vargas, P. (1997). The linguistic intergroup bias as an implicit indicator of prejudice. Journal of Experimental Social Psychology, 33 (5), 490–509.

Appendix

Concept images in experiments 2 and 3

Valence	Type	Descriptor	IAPS Number
Negative	Actions	Heroin	9102
		Attack	6561
		Aimed Gun	6244
		Pipe	2716
		Guns	6830
		Soldier	9160
	Animals	Pitbull	1300
		Dog	1301
		Attack Dog	1525
		Shark	1931
		Leopard	1310
		Shark	1930
	Environment	Puddle	9110
		Oil fires	9120
		Oil fire	9230
		Garbage	9340
		Tornado	5973
		Smoke	9280
	Humans	Angry face	2100
		Skinhead	9800
		Angry face	2110
		Angry face	2120
		Woman	2130
		Boy	2810
	Objects	Aimed Gun	6210
		Bomb	2692
		Tank	6940
		Aimed gun	6190
		Cocaine	9101
		Electric Chair	6020
Neutral	Actions	Smoking	2749

	Casino	7506
	Violinist	5410
	Office	7550
	Chess	2840
	Runner	8465
Animals	Coyote	1640
	Dog	1500
	Dog	1510
	Jaguar	1650
	Grouper	1910
	Fish	1900
Environment	Nature	5220
	Nature	5250
	Plant	5740
	Mushroom	5510
	Winter street	5635
	Skyline	5994
Humans	Neutral face	2200
	Neutral man	2102
	Neutral face	2210
	Man	2190
	Secretary	2383
	Neutral child	2270
Objects	Mug	7009
	Rolling pin	7000
	Mug	7035
	Coffee cup	7057
	Dice	7058
	Book	7090

Attribute words in experiments 2 and 3

Valence	Presented with images of	Word	English meaning
Negative	Actions	förkastlig	reprehensible
		våldsam	violent
		beklaglig	regrettable
		upprörande	outrageous
		stötande	offensive
		oroande	worrying
	Animals	grym	cruel
		hotfull	intimidating
		skräckinjagande	terrifying
		aggressiv	aggressive
		obehaglig	unpleasant
		brutal	brutal
	Environment	förstörd	destroyed
		oroväckande	alarming
		eländig	wretched
		smutsig	dirty
		motbjudande	repulsive
		fördärvad	corrupt
	Humans	ilsken	angry
		provocerande	provocative
		hetsig	heated
		bråkig	rowdy
		arg	angry
		fientlig	hostile
	Objects	farlig	dangerous
		farofylld	dangerous
		skadlig	harmful
		dödlig	fatal
		destruktiv	destructive
		riskabel	risky
Neutral	Actions	ensidig	unilateral
		frekvent	frequent
		generell	general

	indirekt	indirect
	permanent	permanent
	verklig	real
Animals	vaksam	vigilant
	exotisk	exotic
	häpen	amazed
	listig	crafty
	omedveten	unconscious
	utvecklad	developed
Environment	detaljerad	detailed
	evig	eternal
	homogen	homogeneous
	karakteristisk	characteristic
	kortvarig	brief
	enskild	single
Humans	akademisk	academic
	flitig	industrious
	konstnärlig	artistic
	sportig	sporty
	vaken	awake
	vetgirig	curious
Objects	enhetlig	uniform
	praktisk	practical
	beständig	resistant
	enkel	simple
	hållbar	sustainable
	konstgjord	artificial

Attribute words in experiments 1, 4 and 5

Valence	Word	English meaning
Negative	improduktiv	unproductive
	ointelligent	unintelligent
	obegåvad	slow-witted
	fantasilös	unimaginative
	ångestfylld	anxious
	håglös	listless
	inbunden	bound
	inkompetent	incompetent
	opålitlig	unreliable
	ineffektiv	inefficient
	nedstämd	depressed
	otrevlig	nasty
	slarvig	sloppy
	nedslagen	dejected
	trög	sluggish
	passiv	passive
	ängslig	anxious
	olycklig	unfortunate
	ledsen	sorry
	orolig	worried
Positive	skarpsinnig	sagacious
	företagsam	enterprising
	ihärdig	persevering
	handlingskraftig	energetic
	tillförlitlig	reliable
	utåtriktad	outgoing
	älskvärd	amiable
	produktiv	productive
	uthållig	persevering
	målmedveten	purposeful
	energisk	energetic
	dynamisk	dynamic
	ambitiös	ambitious

pålitlig	reliable
kompetent	skilled
begåvad	gifted
skärpt	smart
effektiv	effective
trevlig	nice
aktiv	active