*Implicit Stereotype Content:
Mixed Stereotypes can be Measured with the
Implicit Association Test*

Rickard Carlsson

Magisteruppsats ht 2008

Abstract

According to the Stereotype Content Model (SCM; Fiske, Cuddy, Glick & Xu, 2002), stereotype content can be mixed in terms of diverging evaluation on the warmth and competence dimensions. The present study is the first to demonstrate this with implicit measures. 85 participants completed one of two Implicit Association Tests (IAT; Greenwald, McGhee & Schwartz, 1998), that measured either their associations of lawyers and preschool teachers with warmth or with competence. The participants also completed explicit measures of stereotypes. The results support the SCM, showing that lawyers were implicitly and explicitly associated with less warmth, but with higher competence, relative to preschool teachers. This also illustrates that a single stereotype–IAT may not always be fully representative of stereotype content.

Introduction

Stereotype Content Model (Fiske, Cuddy, Glick & Xu, 2002) postulates that the contents of stereotypes can largely be described as a combination of the two dimensions warmth and competence. According to this model, the contents of stereotypes are sometimes mixed[1] to the extent that some groups (e.g., old people) are perceived as warm, but at the same time as incompetent, while other groups (e.g., rich people) are stereotyped as being cold but still very competent. Indeed, this type of mixed stereotype content have been demonstrated in several studies, encompassing a wide plethora of different samples, including non-student samples (Cuddy, Fiske & Glick, 2007) and cross-cultural comparisons (Cuddy et al., in press).

So far, all studies which have demonstrated mixed stereotype content in terms of diverging warmth and competence, have relied on explicit measures of stereotypes (i.e. self-reports). In fact, there have been very few attempts to systematically study mixed stereotype content with implicit measures (but see Wade & Brewer, 2006). This is somewhat surprising, since implicit measures have mainly two qualities that make them important complements to traditional explicit measures of stereotypes. First, implicit measures are much less susceptible to social desirability concerns (Fazio & Olson, 2003). Second, implicit measures may tap biases that people cannot accurately identify by means of normal introspection and thus could not possibly express explicitly even if they were truly motivated to do so (Greenwald & Banaji, 1995).[2] Hence, implicit measures may prove valuable tools in the study of stereotype content. Consequently, the overall purpose of the present research was to demonstrate mixed stereotype content utilizing implicit measures as a complement to the traditional explicit measures conventionally used in this type of research (e.g., Fiske et al., 2002).

The remainder of the introduction is outlined as follows. The first section deals with theory on stereotype content, and presents the Stereotype Content Model (Fiske et al., 2002) in more detail. The second section reviews current implicit measures with focus on their possibilities and limitations in measuring stereotype content, while the third section deals with the relation between implicit and explicit stereotype content. An overview of the present research concludes the introduction.

---

[1] Henceforth, "mixed stereotypes" will be used to refer to stereotypes with content that can be described as either warm and incompetent, or cold and competent.

[2] Note that while implicit measures have the capability to capture associations that people otherwise could not have accurately identified through introspection (Greenwald & Banaji, 1995), this does not imply that implicit measures always tap associations that are unconscious. Rather, this is an empirical question in each case (see Gawronski, LeBel & Peters (2007) for a review of this topic).

Stereotype Content

In order to make sense of a complex social world without being cognitively overwhelmed, we humans attribute qualities to other individuals based on their social categories (Hilton & von Hippel, 1996). We do this by applying stereotypes, which are "socially shared set of beliefs about traits that are characteristic of members of a social category" (Greenwald & Banaji, 1995:14). The contents of these stereotypes guide what qualities we attribute to the individual, which in turn guide our behavior to the extent that we will act toward the person as if the attributed qualities were true (Greenwald & Banaji, 1995). If stereotypes mostly contained attributes that were favorable (e.g., friendly and intelligent) it would not be very problematic; we would simply have good expectations of people we meet. Alas, stereotypes of out-groups tend to include faulty and unfavorable qualities and thus be a source of prejudice and discrimination (Hilton & von Hippel, 1996).

However, describing stereotypes on a one-dimensional continuum ranging from unfavorable to favorable is an oversimplification, since stereotypes sometimes contain a mix of traits that widely differ in evaluative orientation. Besides several neutral qualities (e.g., rhythmic), the stereotype of a specific group can include beliefs about several traits that are clearly positive (e.g., honest, kind) and at the same time include traits that are clearly negative, such as lazy and unintelligent (Greenwald & Banaji, 1995). In order to describe this mixed content of stereotypes, Fiske et al. (2002) developed the Stereotype Content Model (SCM).

The principle behind SCM is that is that the two dimensions warmth and competence can accurately represent most contents of evaluative stereotypes[3] (Fiske et al., 2002). The warmth and competence dimensions are not unique to the SCM. On the contrary, researchers have used these two dimensions to describe social judgments for decades. Although the definitions and names of the dimensions have varied somewhat, the underlying assumption is that warmth and competence are universal and fundamental dimensions of social judgment, that serve evolutionary purposes (See Fiske, Cuddy & Glick, 2007 for a review).

When humans perceive other humans, it is crucial to determine the other person's intent toward one-self. This corresponds to the warmth dimension, which essentially answers the question regarding whether the other person is a friend or foe. The next step is to gauge the person's capabilities to carry out their intentions, which corresponds to the competence

---

[3] It is important to note that the SCM is a model of evaluative stereotype content. Thus, this model may not fully capture stereotype content that is truly neutral. Still, most stereotype content of out-groups are found to be evaluative, and it is this type of content that are most relevant for prejudice and discrimination (Fiske et al., 2002).

dimension. Although the warmth dimension is primary, and thus to a larger extent influences evaluation as well as behavioral tendencies, both dimensions are important in social judgments. For example, people who have ill intentions toward oneself, but are skilful and competent could prove dangerous opponents, while those with ill intentions but with low competence may not be more than a nuisance (Fiske et al., 2007).

SCM (Fiske et al., 2002) postulates that these basic and universal dimensions of social judgment guide stereotype content. The stereotypes of certain groups include that they are both warm and competent (which is typical of in-groups), while other groups are stereotyped as neither warm nor competent (e.g., homeless people). Stereotypes with these types of contents closely correspond to a one-dimensional good-bad dimension. However, the stereotype content of other groups is mixed. Some groups are stereotyped as warm but still incompetent (e.g., old people), and other groups (e.g., rich people) are perceived as highly competent but cold. These differences in stereotype content are systematic and can be predicted by intergroup relations such as status and competition (Fiske et al., 2002).

Furthermore, there is some evidence that the four specific combinations (e.g., competent but not warm) are qualitatively different in that they evoke four distinct prejudiced emotions (Fiske et al., 2002), which in turn are related to different types of behavioral (i.e. discrimination) tendencies (Cuddy et al., 2007). For example, groups that are perceived as both very incompetent and very cold (e.g., homeless people) evoke disgust (Fiske et al., 2002), are more likely to be the target of passive (e.g., neglect) as well as active harm (e.g., attack) (Cuddy et al., 2007) and the mere perception of them may initiate neurological response consistent with dehumanization (Harris & Fiske, 2006). In contrast, a group that is perceived as warm but not competent (e.g., old people) will evoke pity (Fiske et al., 2002), and while such groups may still be targets of passive harm (i.e. neglect), they may at the same type be the target of active help rather than active harm (Cuddy et al., 2007). In other words, members of such groups may be liked but still disrespected (Fiske et al., 2002).

Implicit Measures of Stereotypes

The advent of implicit measures has truly revolutionized the study of stereotypes and attitudes. What started with a few pioneering studies in the 1980s (e.g., Devine, 1989), became common practice in research following the introduction of the Implicit Association Test (Greenwald, McGhee & Schwartz, 1998). Utilizing the IAT, researchers have demonstrated that negative implicit attitudes and stereotypes of social groups are prevalent and exist among all types of people (Nosek et al., 2007), and that these biases can predict

intergroup discrimination (see metaanalysis by Greenwald, Poehlman, Uhlmann & Banaji, in press). Compared to other implicit measures, the IAT produces large and reliable effects. Moreover, numerous studies have investigated and confirmed the validity of the IAT (see Lane, Banaji, Nosek & Greenwald, 2007, and Fazio & Olson, 2003, for reviews).[4]

The IAT is a computer based experimental paradigm that can detect automatic associations between concepts and attributes. Currently there are four types of IATs commonly used in research. Two of these IAT types are self-esteem and self-concept IATs and will not be described in further detail here. The other two types of IATs are the attitude and the stereotype IAT (Greenwald et al., in press).

The attitude IAT involves pairing two social groups (e.g., Black people and White people) with two valence attributes (e.g., Good, Bad). In contrast, stereotype IATs involve pairing two social groups (e.g., Women and Men) with two attributes (e.g., Family and Career) that have some meaning besides valence (Lane et al., 2007). In some studies, these attributes are neutral or matched in valence to ensure that the IAT does not measure evaluative associations at all (e.g., Amodio & Devine, 2006). Yet, in some studies (e.g., Schwartz, Vartanian, Nosek & Brownell, 2006; Rudman & Lee, 2002; Rudman & Ashmore, 2007) the stereotypes investigated have a clear evaluative component (e.g., lazy, criminal, honest). Thus, these studies have investigated so-called evaluative stereotypes (Rudman & Lee, 2002).

So far, researchers have successfully used implicit measures in studying neutral as well as evaluative stereotypes. However, no studies have demonstrated mixed content in the studied implicit stereotypes. In fact, few studies even discuss to which extent the associations measured are representative of the (hypothesized) stereotype content (but see Amodio & Devine, 2006). This is could be problematic, since IATs are always confirmatory of previous hypothesized associations between concepts and attributes, rather than exploratory. Thus, assuming that the predications of SCM are valid, an evaluative stereotype IAT may indicate negative, positive or neutral associations toward the very same group, depending on whether it captures only the warmth dimensions, only the competence dimension, or some combination of the two dimensions. If this were true, then measuring implicit evaluative stereotypes on a single dimension—as is conventional at this point—would, if the group's stereotype content were indeed mixed, yield an incomplete or even misleading estimate of

---

[4] It should be emphasized that the validity and reliability of the IAT is entirely dependent on that the research question at hand adheres to the methodological limitations of the IAT, for example the relative nature of the IAT. If this is not the case, other implicit measures (see Fazio & Olson, 2003 for an overview) will surely surpass the IATs usefulness.

people's implicit stereotypes. To investigate this possibility, it is important to independently capture both the warmth and the competence dimensions with implicit measures.

Relation between Explicit and Implicit Stereotype Content

Whether the mixed content found on the explicit level is likely to replicate on the implicit level is not entirely clear. Up to this point, there has only been one attempt to disentangle warmth from competence using implicit measures. Wade and Brewer (2006) investigated implicit stereotype content toward two subgroups of women (Businesswomen and Homemakers). Although, the stereotype content of the groups was found to be mixed on the explicit level, the implicit measure (a Lexical Decision Task) did not reveal this pattern. Rather, their results suggest that the implicit measure only tapped evaluative associations (i.e. good-bad).

Thus, the only study of implicit stereotype content so far suggests that mixed stereotype content does not exist on the implicit level. However, there are several reasons to study this question further. First, the reason that the implicit measure used by Wade and Brewer (2006) only captured evaluation may have been method–specific and it is possible that an IAT would yield different results. It is also possible that Wade and Brewer (2006) did not find any mixed stereotypes on the implicit level because of the groups they investigated. In fact, while some studies have, as Wade and Brewer did, found the explicit stereotype content of homemakers to be mixed (e.g., Fiske et al., 2002), other studies have found the explicit stereotype content of the group to be univalent positive – that is both warm and competent (e.g., Cuddy et al. 2007).

Even if the stereotypes of the groups investigated are clearly mixed on the explicit level, and the implicit measures employed are sensitive enough to pick up difference in warmth and competence, the results of implicit and explicit measures may differ for several other reasons. That is, some groups found to be stereotyped differently in terms of warmth and competence on the explicit level might not be so on the implicit level, and the other way around. In fact, the prime reason for using implicit measures in the first place is that implicit measures could yield results that are distinct from results achieved through explicit measures, since implicit measures tap biases that people are sometimes not willing, or able to express in self-reports (Greenwald & Banaji, 1995; Fazio & Olson, 2003). For example, people may report that a specific group is warm but incompetent because it is socially sensitive to state that the particular group is cold, while it is not as sensitive to state that the group is incompetent. Moreover, people may not be fully aware that they have implicit associations that the

particular group is cold (e.g., unfriendly, dishonest) and even if they are aware of it, they may not report it because they do not endorse such views.

This possible discrepancy presents certain challenges in validating implicit measures of stereotype content, since results that conflict with explicit measures will be hard to interpret. For example, if a stereotype that is found to be mixed on the explicit level, is found to be univalent on the implicit level (as in Wader & Brewer, 2006), does this imply methodological problems or actual differences between the explicit and implicit constructs? Such question are, unfortunately, hard to answer prior to rigorous validation of implicit measures of stereotype content. Hence, it is initially important to study groups that can be assumed to show similar results on both explicit and implicit measures.

For long, researchers hypothesized and found little or no relationship between implicit and explicit measures. However, recent research suggests that these disassociations could have been due to a number of reasons including methodological issues such as low reliability in some implicit measures or low correspondence in the design of implicit and explicit measures. For example, if the implicit measure captures relative differences between two groups while the explicit measures captures the absolute evaluation of a single group, the two measures might be disassociated even if the actual underlying constructs indeed are related (Hofmann, Gawronski, Gschwendner, Le & Schmitt, 2005). Another strong moderator of the relationship between explicit and implicit measures is how socially sensitive the topic at hand is. The more sensitive topic, the less negative bias people express in explicit self-reports. This in turn lowers the correlation with the implicit measure (Nosek 2005; Hofmann et al., 2005)[5]

As concern the IAT, a metaanalysis by Hofmann et al. (2005) suggest that there is on average a weak, or close to moderate, relationship with explicit measures. Moreover, Nosek et al. (2007) present results from a study of 17 topics, where over two million IAT sessions and self-reports were correlated. They found that the correlation between IAT and explicit measures (including both attitude and stereotype-IATs) is barely moderate on average, but that there is great variation in this strength depending on the topic. Furthermore, explicit and implicit stereotypes were on average only weakly correlated. However, the correlations were positive for all topics, and the results of both the IAT and the explicit measures were always clearly in the same direction. For example, for groups that were associated with negative

---

[5] According to Nosek (2005) several other variables moderate the relationship as well, such as evaluative strength. Associations that are strong and consistent on both the explicit and implicit level are also more highly correlated. This explains why certain topics have low correlations although they are not socially sensitive (e.g., attitudes toward certain types of clothes).

attributes on the implicit level, the participants also explicitly stated negative views about the group.

In sum, if the topic at hand is not socially sensitive and there is good methodological correspondence between the implicit and explicit measure, we can expect implicit and explicit stereotypes to be slightly related, and at the very least display effects that are in the same direction.

The Present Research

The purpose of the present study was to investigate whether the two[6] mixed combinations of stereotype content postulated by SCM (Fiske et al., 2002), often found on the explicit level also appear on the implicit level. Specifically, to, by the means of the IAT, investigate if the stereotype content of one group includes that the group is warm but incompetent, and the stereotype content of another group includes that that group is cold but competent. To this end, two different implicit measures, one warmth-IAT and one competence-IAT were developed. Furthermore, this study also aimed to examine the relation between these two implicit measures, and their (traditional) explicit counter-parts.

The first step of the present research was choosing which groups to include. It was crucial to choose two groups where the explicit stereotype content is clearly mixed in each own direction. That is, one group stereotyped as warm and incompetent and one group stereotyped as cold and competent. It was also important to choose groups that people do not consider it to be socially sensitive to express negative views about, since previous research (e.g., Nosek , 2005) has found that the relation between explicit and implicit measures is moderated by how socially sensitive the topic under investigation is. Furthermore, since the IAT requires associations to be measured relatively between two groups, it was also important that these groups were entirely distinct and possible to directly compare[7]. In order to identify two groups that met these criteria a pre-study was carried out.

---

[6] While SCM postulates four combinations, only two are mixed. Because the two univalent positive and negative forms of stereotype content have to some extent already been investigated elsewhere (e.g., Rudman & Ashmore, 2007), this fell outside the scope of the present study.

[7] For example, old people cannot be contrasted with rich people, since it entirely possible to be both old and rich.

Pre-study

Method

   Participants and procedure. Fifty-seven students (34 women, 23 men) from various faculties were located at the campus library and asked to fill out a survey about how people in general view social groups. Mean age was 23.77 (*SD* = 4.33).

   Materials. The survey consisted of eleven social groups to evaluate. Five of these groups were chosen because previous research (Cuddy et al., 2007; Fiske et al., 2002) have found that their stereotypes are clearly mixed in that they are warm and incompetent (old people, housewives[8], people with mental retardation) or cold and competent (Asians, rich people, businesswomen). The survey also included profession-groups that were assumed not be socially sensitive to express negative opinions about. This included one group believed to be stereotyped as warm but not so competent (preschool teachers) and two groups (mathematicians, lawyers) believed to be perceived as cold but very competent. Finally, two groups that could be assumed to be perceived as low in both competence and warmth were included for calibration (immigrants, criminals).

   The participants evaluated the eleven groups on a competence and on a warmth scale, see Table 1. The scales were Swedish adaptations of the scales previously used in research of stereotype content (e.g., Fiske et al., 2002, Cuddy et al., in press), and each contained four items. The participants made the ratings on a scale ranging from 1 (not all) to 5 (extremely) regarding how the group is viewed in Swedish society (e.g., extremely honest). As conventional (Fiske et al., 2002), the participants were instructed not to give their own opinion, but rather the views of people in general. The questions were framed this way in order to reduce social sensitivity[9]. For each group, the participants also rated how socially sensitive it is to express negative views about the group. The Swedish version of the scales can be provided upon request.

---

[8] The survey used the term housewives rather than the preferred (more inclusive and less pejorative) term homemakers, since the latter have no Swedish equivalent that is widespread enough to be recognized by the participants in the present study.
[9] Social desirability concerns are also the reason why all the items were framed as positive traits. It is arguably less socially sensitive to state that a group is not at all intelligent, than to state that the group is extremely unintelligent (see Fiske et al., 2002).

Table 1

The scales and their respective items

Scales	Items
Competence	As viewed by society how […] are members of this group? [competent, intelligent, skillful, capable]
Warmth	As viewed by society how […] are members of this group? [warm, friendly, honest, well-intentioned)
Social sensitivity	How sensitive is it to express negative views about this group?

Results and discussion. The reliability of the scales was satisfactory with Cronbach's alpha > .85 for all of the eleven groups and both scales. Paired-sample t-tests were used to test for differences in ratings of the competence dimension and the warmth-dimension for each group. For all groups there were significant differences in the rating of competence and warmth, suggesting mixed stereotypes ($p < .01$ for all comparisons, except for Asians where $p = .05$). Importantly these differences were in the expected directions for all groups. For example, people with mental retardation were rated as less competent than they were rated as warm, while rich people were rated as more competent than warm, see Table 2.

Somewhat surprisingly, the two groups (immigrants, criminals) included that were assumed to be univalent negative were also found to have mixed stereotypes. However, a closer look at the means (see Table 2) reveals that although criminals were perceived as more competent than they were perceived as warm both ratings were very low compared to the other groups. The results show a similar pattern for immigrants.

Finally, as expected, the participants reported that it was more sensitive to report negative views about some groups then of others, with certain groups being quite sensitive (e.g., people with mental retardation) to express negative opinions about and other much less so (e.g., lawyers), see Table 2.

Table 2

Warmth, competence and social sensitivity rating of the groups in the survey

Social group	Warmth		Competence		Social sensitivity	
	Mean	SD	Mean	SD	Mean	SD
Asians	3.45	0.80	3.66	0.75	3.00	1.16
Businesswomen	2.99	0.82	4.04	0.68	2.58	1.16
Criminals	1.37	0.57	2.50	0.97	1.65	0.97
Housewives	3.85	0.79	2.73	0.76	2.65	1.06
Immigrants	2.39	0.75	2.67	0.68	3.61	1.10
Lawyers	2.46	0.83	4.23	0.60	1.74	0.94
Mathematicians	2.96	0.53	4.53	0.62	1.75	0.89
Old people	3.81	0.73	3.23	0.64	2.88	1.12
People with mental retardation	3.84	0.88	2.06	0.70	4.18	1.07
Preschool teachers	4.23	0.63	2.93	0.68	2.56	1.10
Rich people	2.25	0.67	3.51	0.72	2.00	1.05

Note: A higher number indicates that the participants rated the group as warmer, more competent or more sensitive to express negative views about. The differences in warmth and competence rating are statistically significant for all groups ($p < .01$), except for Asians ($p = .05$).

Based on the result of this pre-study, two groups were chosen to be included in the main study: lawyers and preschool teachers. The stereotype content of these groups were clearly mixed in different directions, with lawyers perceived as competent but not warm, and preschool teachers as warm but not competent. Moreover, paired sample t-tests confirmed that lawyers were perceived as more competent ($M = 1.30$, $t(56) = 10.09$, $p < .01$) , but at the same time as less warm ($M = -1.78$, $t(56) = -13.19$, $p < .01$), when compared to preschool teachers.

Main Study

Guided by the results from the pre-study, the following predictions were made. First, it was predicted that lawyers would be associated with less warmth relative to preschool teachers on both the implicit and the explicit level. Furthermore, lawyers were predicted to be associated with higher competence relative to preschool teachers, again on both the implicit and the explicit level.

Regarding the relationship between the implicit and the explicit measures, it was predicted that the implicit measure of warmth would be significantly and positively correlated with the explicit measure of warmth, but be unrelated to the explicit measure of competence. Likewise, it was predicted that the implicit measure of competence would be positively related to the explicit measure of competence, but not to the explicit measure of warmth.

Method

Participants. The participants were recruited in the same way as in the pre-study, and were thus volunteering students from various faculties. Five participants did not complete the whole procedure because of computer malfunction. In the end, the participants were 85 (53 women, and 32 men). Mean age was 23.29 (*SD* = 4.23).

Explicit measures. The explicit measures were constructed with respect to two primary aims. First, they had to be closely similar to explicit measures in the pre-study, and thus similar to measures used in previous research on the SCM (e.g., Fiske et al., 2002). Second, in order to maximize correlations with the implicit measures they had to be similar to the implicit measures in construction. Consequently, the measures were based on previous measures that have been found to correlated with the IAT (Nosek et al., 2007), but involved the same traits that the items did in the pre-study. The participants were instructed to for each trait (four warmth and four competence traits) rate how strongly they associated the trait with lawyers and preschool teachers. There were seven alternatives ranging from "Much stronger with lawyers" (coded as 3) to "Much stronger with preschool teachers) (coded as -3), with the neutral intermediate (coded as 0) "Equally strong with lawyers and preschool teachers".

Implicit measures. The warmth-IAT had the attribute labels "Warm" and "Cold". The stimuli belonging to the warm category were four words identical to the traits that comprised the explicit warmth-scale; the stimuli belonging to the cold category were four words that were approximate opposites of the warmth traits (e.g., dishonest). Similar, the competence-IAT included the attribute labels "Competent" and "Incompetent", with stimuli consisting of the words that made up the explicit competence-scale along with matching opposites. Table 3 summarizes the difference in attribute labels and stimuli for the two IATs.

Table 3

The attribute labels and stimuli of the IATs

IAT	Category and stimuli words
Competence-IAT	Competent: competent, intelligent, skillful, capable
	Incompetent: incompetent, unintelligent, incapable, unqualified
Warmth-IAT	Warm: warm, friendly, honest, well-intentioned
	Cold: cold, dishonest, disagreeable, unfriendly

Except for these differences in attribute labels and stimuli, the two IATs were identical in every detail. The target labels consisted of preschool teachers and lawyers. The corresponding

stimuli for these categories were chosen with careful consideration to avoid any confounds. Photos of preschool teachers and lawyers was not used, since it could be problematic to match these in every possible way, due to for example different clothing typically used and different work situations (e.g., a representative photo of a preschool teacher would include children). Instead, the stimuli consisted of synonyms to the profession titles: lawyer, jurist and legal–representative for the lawyer category, and preschool-teacher, kindergarten-staff and pedagogue for the preschool teacher category.[10] This choice limited the number of stimuli, because there are simply not that many synonyms in Swedish. However, previous studies have shown that the IAT can produce reliable and valid results with as few as two stimuli per category (Nosek, Greenwald & Banaji, 2005).

The design of the IAT followed the recommendations in Nosek et al. (2005) and Lane et al. (2007). Each IAT comprised five parts over a total of seven blocks each consisting of 20–40 trials. Following is a step-by-step description of the warmth-IAT, which also applies for the competence IAT.

In the computerized test, two keys "e" and "i" were mapped to categories appearing at the top left and right side of the screen respectively. Using these two keys, the participants rapidly classified stimuli appearing in the middle of the screen to the correct category. If they made an error, they were prompted to correct the mistake by hitting the other key. For each trial, the response times were recorded from the time the stimulus was presented, until the participant made the correct response.

The first two parts consisted of two blocks where the participants familiarized themselves with the procedure and learned the correct key mappings. In the first part (block 1) the participants classified attribute stimuli (e.g., honest, dishonest) to the two attribute categories (Warm, Cold) for 20 trials. In the second part (block 2), the participants instead classified target stimuli (e.g., jurist, pedagogue) to the two target categories (Lawyer, Preschool-teacher) for 20 trials.

The third part was the first critical part of the IAT, and comprised two blocks where the participants classified both target and attribute stimuli to the four categories labeled in the right and left upper corner of the screen, while still using only two keys. Half of the participants performed the hypothesized compatible part in this third part. Thus, they pressed "e" for stimuli belonging to either "Preschool-teacher" or to "Warm", while they pressed "i"

---

[10] Note that since the IATs were presented in Swedish, and the stimuli presented here are direct translations, the choice of stimulus words may not be the words best suited for an English version of the IAT. The full list of the Swedish words used can be found in the appendix.

for stimuli belonging to either "Lawyer" or "Cold". The other half of the participants performed the incompatible part (Lawyer + Warm / Preschool-teacher + Cold). Counterbalancing the compatible and incompatible parts in this manner is standard procedure, because it reduces the risk of an order effect that sometimes slightly affects the magnitude of the IAT results in that performing the compatible part first yields stronger effects than performing the incompatible part first (Nosek et al., 2005). The participants classified stimuli in this manner for 20 trials in block three, and for 40 trials in block four.

In the fourth part (block 5), the two target categories switched key mappings with each other. As in part two, the participants only classified target stimuli to the two target categories, but this time for 40 trials in order to learn this new key mapping properly.

The fifth part (block 6, 7) was identical to the third part with the one exception that the participants now classified stimuli with the new key mapping learned in the fourth part. Hence, participants that had the compatible key mapping (Preschool-teacher + Warm / Lawyer + Cold) in part three now had the incompatible key mapping, and the other way around.

The IAT is in it self a repeated measure experiment, since each participant performs both the incompatible and compatible combination that differs only in that key combinations differ. The difference in response times for these two parts is the central idea of the IAT. A shorter response time for a combination indicates that this better fits the participants' associations. For example, if the compatible part in the warmth-IAT is performed faster than the incompatible part, this indicates that the participants implicitly associate preschool teachers with warm and lawyers with cold, more strongly than they associate preschool teachers with cold and lawyers with warm.

However, because response times are not normally distributed, and because overall response times are highly individual, simple mean differences in the response between the compatible and incompatible combination may be a poor estimate of the participants' associations. Moreover, individual differences in task switching (i.e. alternating between classifying attribute and target stimuli) may affect the response times, as may extraneous influences. To address these issues, Greenwald, Nosek and Banaji (2003) have developed a scoring algorithm for the IAT, which is now standard in calculating IAT-effects (e.g., Lane et al., 2007).

Following the guidelines of Greenwald et al. (2003), all trials from the two critical parts three (block 3, 4) and five (block 6, 7) were included in the analysis. First the data was screened for suspiciously short response times indicating the participant had pressed the keys

at random. Participants with more than 10% of trials <300ms were removed (however no participants in this study fell in this category).Then all trials above 10 000 ms were removed from the analysis, since such long response are not likely to be valid responses (e.g., somebody coughed).[11] Following this, mean difference in response times for the trials in block 3, and 6 were calculated for each participant. This was then divided by the participant's inclusive standard deviation in response times for these two blocks. The same was done for block 4 and 7. These two figures were averaged, which produced a *D*-value.

The *D*-value is the individual's standardized score of the IAT (Greenwald et al., 2003). For both IATs, the *D*-values were calculated so that a positive value reflected an association of lawyers and favorable attributes (Warm and Competent respectively) relative to preschool teachers. A negative score indicated a more favorable association of preschool teachers relative to lawyers. A value of zero indicated no differences in association.

Procedure. The participants completed the procedure either on a home computer, or on campus, through an online web-version of Inquist (Inquisit, 2008). This software accurately records response times locally on the participant's computer using DirectX-technology and sends the results to a web server through a secure connection. Before launching the applications, the software checks for compatibility with the computer. Thus, internet connection quality, speed of computer etc. should not affect the results. Moreover, the participants received instructions that emphasized that they had to take the test in a calm setting, and to report any problems (e.g., computer malfunctions, or being disturbed during the test) to the researcher.

The participants were randomly assigned to complete either the warmth-IAT or the competence-IAT. The order of which the participants performed compatible and incompatible combinations of the IATs was counter-balanced between participants. After completing the IAT the participants filled out the explicit measures and provided demographic data. The order of the IAT and the explicit measures was not varied between participants, since previous studies have found little or no effects on this order (Hoffman et al., 2005; Lane et al., 2007).

The participants were informed that the purpose of the study was to investigate associations between social groups, in this case job occupations, and traits. Thus, when

---

[11] Additionally, the data were screened for unusual amount of errors (> 40%, following Nosek et al., 2007), which could indicate that the participant had somehow misunderstood the procedure. No participant had that many errors.

providing informed consent, the participants were fully aware of the purpose of the study, but not of the specific hypotheses.

Results

Implicit measures. Based on previous research and the results of the pre-study it was predicted that lawyers would be less associated with warmth, but more associated with competence, relative to preschool teachers, on the implicit level. To test this, the response times of the compatible and incompatible parts of the warmth- and competence-IATs were compared based on the *D*-statistic (described in the method section). In the warmth-IAT, the participants had shorter response times in the compatible (Preschool-teachers + Warm / Lawyers + Cold) than in the incompatible (Lawyers + Warm / Preschool-teachers + Cold) part: mean $D = -0.47$ ($SD = 0.42$). A one-sample t-test confirmed that this *D*-value was significantly different from zero: $t(41) = -7.16$, $p < .01$. Thus, as predicted, the participants implicitly associated lawyers with less warmth relative to preschool teachers.

Also in the competence-IAT, the response times were shorter in the compatible (Lawyers + Competent / Preschool-teachers + Incompetent) than in the incompatible (Preschool-teachers + Competent / Lawyers + Incompetent) part: mean $D = 0.13$ ($SD = 0.37$). A one-sample t-test confirmed that this *D*-value was significantly different from zero: $t(42) = 2.33$, $p < .05$. Consequently, as predicted, the participants implicitly associated lawyers with higher competence relative to preschool teachers.

In order to rule out possible confounding variables, the results of the implicit measures were checked for order effects as well as sex and age differences. The effects of order of the compatible or incompatible parts were non-significant for the two IATs. Moreover, there were no significant differences with regard to the sex or age (or interaction thereof) of the participants.

Explicit measures. Another prediction of the present research was that the participants would explicitly state that lawyers are less warm, but more competent, when compared to preschool teachers. The results confirmed this. The participants rated lawyers as less warm ($M = -1.59$, $SD = 0.87$), but as more competent: ($M = 1.03$, $SD = 0.95$), compared to preschool-teachers. One-sample t-tests showed that these differences were statistically significant ($t(84) = -16.86$ and 9.92 respectively, $p < .01$). Furthermore, the reliability was acceptable (Cronbach's alpha $> .75$) for both the warmth and the competence scale

Explicit–implicit relationship. A final hypothesis of the present research was that the IATs would be positively correlated with their explicit counterpart, but not with the explicit

measure representing the other dimension. That is, that the warmth-IAT would be positively correlated with the explicit warmth scale but not with the explicit competence scale. Likewise, the competence-IAT was predicted to be positively correlated with the explicit competence scale, but not with the explicit warmth scale.

As predicted, the competence-IAT was significantly correlated with the competence-scale: $r = .32$ ($n = 43$, $p < .05$), but not with the explicit warmth-scale ($n = 43$, $r = -.10$, $p = .52$). However, contrary to the predictions, the warmth-IAT was not significantly correlated with the explicit warmth-scale ($n = 42$, $r = .11$, $p = .48$). Furthermore, the warmth-IAT did not significantly correlate with the explicit competence-scale ($n = 42$, $r = .16$, $p = .31$).[12]

Discussion

The present study is the first that demonstrates mixed stereotype content using implicit measures. By combining two IATs—one measuring the warmth dimension and one the competence dimensions— the present study captured the two dimensions of stereotype content (Fiske et al., 2002), resulting in a more detailed description of implicit stereotype content than a single dimensions alone could have. In the present study, two social groups (lawyers and preschool teachers) were contrasted against each other in these IATs. The results show that lawyers were at the same time implicitly stereotyped as competent and cold, compared to preschool teachers. Because the IAT is a relative measure, it also is equally correct to conclude that preschool teachers were implicitly stereotyped as incompetent and warm relative to lawyers. Thus, the results indicate two different types of mixed implicit stereotypes: warm and incompetent, and cold and competent, as postulated by SCM (Fiske et al., 2002).

Moreover, the results on the explicit level show a similar pattern, with the participants stating that lawyers are more competent, but less warm, relative to preschool teachers. Thus, both at the explicit and the explicit level the stereotypes of lawyers and preschool teachers were mixed and in the same direction.

To some extent the positive relationship between the results on the explicit and implicit measures could also be confirmed, since the competence-IAT and the explicit competence scale correlated significantly. This was, however, not the case for the warmth-IAT and the warmth scale. This result is actually not that surprising, since previous studies have shown

---

[12] To ensure that the correlations of the implicit and explicit measures was not affected by a relationship between the two explicit scales, or by the age or sex of the participants, partial correlations where these variables were controlled for were computed. The correlations remained almost identical in magnitude and significance level, suggesting that the results were not to any large extent affected by these variables.

that explicit and implicit stereotypes are typically only weakly correlated (e.g., Nosek et al., 2007). This effectively means that in order to achieve sufficient power to detect such correlations, a sample size of several hundred participants is required. In light of this, it could be argued that investigating these correlations in smaller studies such as the present study was a fruitless endeavour. Still, in the present study several attempts were made to ensure that at least an average correlation could be expected. First, the present study used explicit measures with good methodological correspondence with the IAT, since they were relative and directly (as in the IAT) contrasted the two groups. Moreover, the groups chosen were (according to the pre-study) not so socially sensitive to express negative views about. Still, it is possible that other unobserved variables could have moderated the relationship between the implicit and explicit warmth measure (see Nosek 2005, for a review of such moderators). This is, however, not possible to determine in the present study, but including controls for such moderating variables might be a successful way for future research to determine if the null correlation is a result of low power, or actual disassociation.

There is one other plausible explanation for the lack of clear correspondence between the implicit and explicit warmth measure: the lack of variation in the explicit measure. In fact, all but one of the participants stated that lawyers were less warm than preschool-teachers. Thus, the lack of correlations could be due to a floor effect. If this is the case, then future studies investigating groups with less clear stereotypic associations, or using samples with greater plurality might find greater implicit and explicit correspondence regarding warmth.

Implications for SCM

The present study contributes to the growing literature supporting the Stereotype Content Model (e.g., Fiske et al., 2002; Cuddy et al., in press). Because previous studies based on the Stereotype Content Model exclusively tested the model with explicit methods, alternative explanations of the results could not be ruled out. For example, as discussed briefly by Fiske et al. (2002), mixed stereotypes could possibly be due to a conflict between egalitarian values and prejudice. That is, although people actually stereotype certain groups as altogether negative, they do not express this explicitly, but rather state some negative opinions about a group and some positive. If this were the case, then mixed stereotypes should primarily show up in self-reports and not in implicit measures that are less affected by social desirability concerns, as well as rely on automatic and spontaneous responses rather than reasoning (Fazio et al., 2003).

However, the results of the present study suggest otherwise. At least for two groups (preschool teachers and lawyers) the predictions of SCM hold up when using implicit measures (IAT). Importantly, both mixed combinations (warm, incompetent; cold, competent) were confirmed on the implicit level in addition to the explicit level. While the present study on its own can only offer some preliminary evidence that stereotype content can be mixed on the implicit level, it does add to the large amount of research suggesting that warmth and competence are important dimensions of social perception (see Fiske et al., 2007).

Implications for IAT Research

Although literally hundreds of studies utilizing the IAT have been published (see Lane et al., 2007 for a review) and many of them have focused on stereotypes (e.g., Scwhartz et al., 2006; Rudman & Ashmore, 2007), there are up to this date no studies that have demonstrated that the stereotypes of the groups investigated are mixed. The results of the present study suggest that stereotype content indeed may be mixed on the implicit level. In doing so, the present study highlights the risks of drawing conclusions based on a single IAT of associations toward a group. For example, from the present study it is possible to draw the conclusion that lawyers are cold and thus associated with negative features. Still lawyers were also found to be competent and thus associated with many positive features (such as intelligence). Drawing conclusions based only on one of these two IATs would not have been fully representative of how the group is associated on the implicit level.

It is possible that the groups investigated in this particular study are a special case in that the stereotypes are mixed and that most groups are not stereotyped in this mixed way on the implicit level. However, studies using explicit methods suggest that mixed stereotypes are very common (Fiske et al., 2002). Moreover, the pre-study in the present paper, found that stereotypes were mixed for 10 out of 11 groups. This point to that future research on stereotype IATs should, at the very least, consider that the specific associations measured may not be fully representative of how the group is associated, and that other IATs (i.e. a warmth instead of a competence-IAT) might produce clearly diverging results.

Limitations of the Present Study

It could be argued that the present study did not capture stereotype content of lawyers and preschool-teachers, but of men and women in general. Indeed, it is possible, and even likely, that preschool teachers are more strongly associated with women, and lawyers more strongly associated with men. However, women and men tend to associate their own gender with

positive traits on implicit measures (Rudman, Greenwald & McGhee, 2001). Thus if the present study mainly captured gender stereotypes, then we would expect that women associated preschool-teachers with both competence and warmth, and that men associated lawyers with both competence and warmth. Yet, there were no such pattern to the results. Moreover, in the present study no masculine or feminine descriptions of either lawyers or preschool teachers were used, nor were any pictures with women or men that could have affected the participants' associations presented. Thus, it was entirely up to the participants how they associated the two groups.

One limitation in the present study is that only two groups were investigated. Clearly, stereotype content of additional groups will have to be investigated before drawing any strong conclusions. This is especially important since the only previous study of mixed implicit stereotypes, did not find mixed stereotypes on the implicit level (Wade & Brewer, 2006).

It is also important to note that since the IAT used in the present research, only captures relative differences, the result tells us nothing about how preschool teachers or lawyers are associated on an implicit level in an absolute sense. For example, it is not possible to conclude whether it is that lawyers are associated as particularly competent or that preschool-teachers are associated as particularly incompetent; it is only possible to conclude that there is a difference. Thus, the practical application of the present studies results are limited to the situations were lawyers and preschool teachers are contrasted against each other. Still, these relative differences do reflect something about how these two groups are viewed in society.

Future Research

The only previous study investigating mixed stereotypes with implicit methods (Wade & Brewer, 2006) suggest that this does not exist. Thus, at this point the only two studies of this subject show conflicting results. Hence, future studies are important. A natural next step is to study the same groups as Wader and Brewer (2006) did (homemakers, businesswomen) but with the methodology used in the present study (i.e. the IAT). If these groups show mixed stereotypes on the implicit level with the methodological approach in the present study, it would suggest that the methods employed by Wader and Brewer simply were not sensitive enough to pick up differences in competence and warmth.

The groups studied in the present study are contrasted against each other, producing relative effects. However, many groups have no clear contrast. For example, contrasting rich people (that were found to be mixed on the explicit level in the pre-study) with poor people would be logically correct. Yet, this would likely be insufficient in detecting mixed

stereotypes, since previous research using non-relative explicit measures (e.g., Fiske et al., 2002) have found poor people to be stereotyped as both incompetent and cold. Thus, rich people, although in absolute terms cold and competent, may be perceived as equally warm (or warmer) as well as more competent relative to poor people. Hence, contrasting two groups in a relative measure is not suitable in all cases. Future research will need to address this with non-relative implicit measures, which although lower in reliability than the IAT, could, if used correctly, improve the understanding of these concepts (see Fazio et al., 2003 for a review of such measures).

Conclusion

It is often both easy and tempting to view complex phenomena on a single bipolar dimension. However, things in life are rarely altogether black or white; nor are multiple shades of grey sufficient to explain multifaceted issues such as inter-group bias. Still, research into prejudice and stereotypes traditionally focuses on whether a group is perceived as good, bad or somewhere in between (Fiske et al., 2002). Several studies using explicit methods (e.g., Cuddy et al., 2007, Fiske et al., 2002) have already demonstrated that this classification is an oversimplification and that describing stereotypes in terms of the two dimensions warmth and competence provide a more accurate and detailed description of stereotype content. The present study add to this line of research by demonstrating that stereotypes are sometimes mixed in this manner on the implicit level as well. Thus, even the automatic and spontaneous components of inter-group bias may sometimes be too complex to be described as simply good or bad.

References

Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: Evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology, 91*(4), 652-661.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology, 92*(4), 631-648.

Cuddy, A. J. C., Fiske, S. T., Kwan, V. S. Y., Glick, P, Demoulin, S., Leyens, J. Ph., Bond, M. H. et al. (in press). Stereotype content model holds across cultures: Universal similarities and some differences. *British Journal of Social Psychology.*

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*(1), 5-18.

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

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77-83.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878-902.

Gawronski, B., LeBel, E. P., & Peters, K. R. (2007) What do implicit measures tell us? Scrutinizing the validity of three common assumptions. *Perspectives on Psychological Science, 2(2)*, 181-193.

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

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. (2003). Understanding and using the implicit association test: I. an improved scoring algorithm. *Journal of Personality and Social Psychology, 85*(2), 197-216.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E., & Banaji, M. R. (in press). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology.*

Inquisit 3.0.3.0 [Computer software]. (2008). Seattle, WA: Millisecond Software.

Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the lowest of the low: Neuroimaging responses to extreme out-groups. *Psychological Science, 17*(10), 847-853.

Hilton, J. L., & von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology, 47*, 237-271.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31*(10), 1369-1385.

Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and using the implicit association test: IV: What we know (so far) about the method. In B. Wittenbrink, & N. Schwarz (Eds.), *Implicit measures of attitudes.* (pp. 59-102). New York: Guilford Press.

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General, 134*(4), 565-584.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the implicit association test: II. method variables and construct validity. *Personality and Social Psychology Bulletin, 31*(2), 166-180.

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., Smith, C. T., Olson, K. R., Chugh, D., Greenwald, A. G., & Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology*, *18*, 36-88.

Rudman, L. A., & Ashmore, R. D. (2007). Discrimination and the implicit association test. *Group Processes & Intergroup Relations, 10*(3), 359-372.

Rudman, L. A, Greenwald, A. G., & McGhee, D. E. (2001). Implicit self-Concept and evaluative implicit gender stereotypes: Self and ingroup share desirable traits. *Personality and Social Psychology Bulletin*, 27(9), 1164-1178.

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

Schwartz, M. B., Vartanian, L. R, Nosek, B. A., & Brownell K. D. (2006). The Influence of one's own body weight on implicit and explicit anti-fat bias.*Obesity, 14*, 440-447.

Wade, M. L., & Brewer, M. B. (2006). The structure of female subgroups: an exploration of ambivalent stereotypes. *Sex Roles: A Journal of Research, 54*, 753-765.

Appendix

Table 4

The labels and stimuli of the IATs, as presented in Swedish

IAT	Category and stimuli words
Competence-IAT	Kompetent: kompetent, intelligent, skicklig, begåvad
	Inkompetent: inkompetent, trög, oduglig, okunnig
Warmth-IAT	Varm: varm, vänlig, ärlig, välmenande
	Kylig: kylig, ovänlig, falsk, otrevlig
Both IATs	Advokat: advokat, jurist, rättsombud
	Förskollärare: förskollärare, dagispersonal, pedagog