



HAL
open science

The Language of Bias: Unveiling Implicit Attitudes through Linguistic Markers in Maghrebians Prejudice

Brigitte Bardin, Anne Burguet, Sophie Henry, Nicolas Souchon

► **To cite this version:**

Brigitte Bardin, Anne Burguet, Sophie Henry, Nicolas Souchon. The Language of Bias: Unveiling Implicit Attitudes through Linguistic Markers in Maghrebians Prejudice. 2023. hal-04366816

HAL Id: hal-04366816

<https://hal.science/hal-04366816>

Preprint submitted on 29 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

This is a preprint. The paper is currently undergoing peer-review and is subject to change.

The Language of Bias: Unveiling Implicit Attitudes through Linguistic Markers in Maghrebians Prejudice

Brigitte Bardin (1), Annette Burguet (2), Sophie Henry (1), Nicolas Souchon (3)

(1) Parisian Laboratory of Social Psychology (LAPPS), ENOSIS, University Paris 8 Vincennes, Saint Denis, France.

(2) Laboratory of Applied Studies and Research in Social Sciences (LERASS), PsyCom, University Toulouse 3, France.

(3) Laboratory on Interactions Cognition, Action, Emotion (LICAE), University Paris 10, Nanterre, France.

Corresponding author: Brigitte Bardin, Laboratoire Parisien de Psychologie Sociale (LAPPS), 2 rue de la Liberté, 93526 Saint-Denis cedex, France. brigitte.bardin@univ-paris8.fr

The Language of Bias: Unveiling Implicit Attitudes through Linguistic Markers in Maghrebians Prejudice

Abstract

The challenge of measuring prejudice without it being influenced by social desirability has prompted the development of indirect measures of attitude. The effectiveness of implicit association tests (IAT, SC-IAT-P) in assessing racial prejudices and their complementarity with self-reported measures has been demonstrated. Additionally, research on intergroup linguistic biases (LIB) has revealed that individuals tend to employ more abstract language in negative discourse about an outgroup and more concrete language in positive discourse when expressing prejudices. Furthermore, individuals who are sensitive to social desirability, characterized by a relatively high level of education and/or a left-leaning political orientation (as opposed to right-leaning), would be likely to exhibit a higher level of language abstraction when expressing prejudice towards an outgroup.

The objective of this series of studies is to evaluate the relationship between two indirect measures of prejudices towards Maghrebians (LIB and SC-IAT-P) and an explicit measure of prejudice, while examining the moderating effects of social desirability and political orientation. Since Study 1 did not provide conclusive evidence of a correlation between these two indirect measures, methodological biases were addressed, and social desirability and political orientation were introduced as moderators in Study 2. Consequently, a moderated moderation model (political orientation*social desirability) was employed, revealing a significant moderate moderation effect only for the links between SC-IAT-P and explicit prejudice but not for LIB as antecedent. This finding led to further analysis and the assessment of a mediation model in which LIB mediated the relationship between the SC-IAT-P measure and explicit prejudices towards Maghrebians. The main results support this model. Beyond confirming the relevance of an indirect measure of prejudice through an implicit association test (SC-IAT-P), the results also affirm the utility of LIB

as a linguistic indirect measure of prejudice. Therefore, analyzing the degree of linguistic abstraction in any discourse could be a potential indicator of the valence of speakers' attitudes towards an outgroup.

Key words

Indirect measures of attitude, LIB, SC-IAT-IP, prejudice, political orientation, social desirability

The authors declare no conflicts of interest. Ethical approval was obtained for this research by the ethics committee of the University of Toulouse - France (reference: 2020-295). The data and materials used in this study are freely accessible: <https://researchbox.org/>

The Language of Bias: Unveiling Implicit Attitudes through Linguistic Markers in Maghrebians Prejudice

Assessing prejudice is an essential part of psychosocial research. Self-reported measures are the most widely used to account for attitude towards a stereotyped group. However, their use reveals two difficulties particularly. Thus, people demonstrate social desirability and adapt the expression of their attitude to social norms in order to present a positive image (Campbell & Stanley, 1963) particularly when attitudes towards groups which are subject of discrimination (such as racial prejudices, gender prejudices, or addictive behaviors). Use of self-reported measures also raises questions about individuals' conscious accessibility to their own attitude through introspection (Nisbett & Bellows, 1977; Nisbett & Wilson, 1977). For the past 30 years, many researchers have been investigating new measures related to the automatic triggering of an evaluation of the object at its mere evocation (Bargh et al, 1992; Fazio et al.,1986). Research carried out in the field of implicit social cognition attest to the predictivity of this implicit attitude on behavior (Perugini et al., 2010). The growing interest in this field of research has led to the development of new attitude measurement tools and the testing of their psychometric characteristics. The IAT (Implicit Association Test; Greenwald et al., 1998) is the most popular tool and has been used in over 550 studies since 1998 in various fields, including racial prejudices (e.g. Greenwald et al., 1998; Rudman et al., 1999; Dasgupta et al., 2000; McConnell & Leibold, 2001; Gawronski, 2002). The IAT is based on reaction times to categorize two target concepts (e.g. female *versus* male) in association with two evaluative concepts (e.g. bad *versus* good). IAT is an informatized or pen & pencil measure. It measures reaction time at which participants associate concepts with categories and is done by comparing cross associations (e.g., concepts "female" and "good" *versus* "male" and "good" *versus* "female" and "bad" *versus* "male" and "good"). However, two main criticisms are made about it. On the one hand, it is a relative measure, as it measures the

preference of one object over another. On the other hand, its measurement is contaminated by extra-personal associations, that is, by social norms (e.g., Azar, 2008; Oswald et al., 2013 ; Kaskeleviciute et al.,2023). To overcome these two main limitations, variants of the IAT have been developed, such as the SC-IAT-P (Bardin et al., 2014). The SC-IAT-P measures implicit attitude towards a single object, rather than comparing associations between different social groups. It also limits the contamination of its measurement by extra-personal associations. While Greenwald et al.(1998) argue that the IAT measures individual differences, Schimmack (2019) argues that there is insufficient evidence to support this position. According to this author (2019), only 20% of the variance of the IAT as a measure of attitude towards an ethnic origin reflects racial preference. Von Hippel et al. (1997) justify the various results regarding the different levels of correlation between implicit measures (IAT) and explicit measures (Fazio et al., 1995; Wittenbrink et al., 1997) by the nature of the explicit measure, which can more or less predict biases depending on the level of social desirability felt by participants, the difficulty of consciously accessing biases, or a context inducing a politically correct attitude. However, if these factors are not salient, self-reported measures could be equally valid in predicting detrimental attitudes as implicit measures. However, Schimmak (2019) notes that the IAT can be a relevant complementary measure in the study of constructs that are sensitive to social desirability. Overall, the author highlights a lack of construct validity in the use of the IAT. The construct validity of an assessment tool can be evaluated by comparing the results obtained using different tools that are supposed to measure the same construct. Therefore, it is not about repeating the same IAT measure over time, but rather using different, implicit or explicit, measures of attitude towards the same object.

Following Von Hippel et al.,(1997) degrees of language abstraction is an implicit marker of prejudice. Maass et al.,(1995) found that the degree of language abstraction is different depending on whether people describe stereotypical or non-stereotypical behaviors of an out-group. Thus, stereotypical behaviors will be described in more abstract terms than non-stereotypical behaviors

(Linguistic Intergroup Bias, LIB, Maass et al, 1999 ; Menegatti & Rubini, 2009; Wigboldus & Douglas, 2007 ; Wigboldus et al., 2000), all the more so because the level of prejudice of the describer is high (Schnake & Ruscher, 1998) and their degree of social proximity with the target to describe is low (Li & Hills, 2021 ; Snefjella & Kuperman, 2015). Concrete terms are considered to describe an unstable characteristic because they specify the behavior in a context, thereby making it subject to that context. They are considered less general and less likely to be used across different contexts. In contrast, abstract terms represent more stable characteristics because they refer to more general concepts or attitudes. The previous research (e.g., Dragojevic et al., 2017 ; Mastro et al., 2014) on Linguistic Intergroup Bias (LIB) found that people communicate about positive in-group behaviors more abstractly than positive out-group behaviors, while negative in-group behaviors are expressed more concretely than negative out-group ones. The subdivision of concreteness and abstraction was initially proposed by the Linguistic Category Model (LCM : Semin & Fiedler, 1988, 1991; see Wigboldus & Douglas, 2007 for a review) as a tool to analyze how individuals use interpersonal terms to describe social events (Semin, 2012) and to communicate their stereotypes (Maass et al., 1989). The authors classify language into four levels going from the most concrete to the most abstract. Thus, the descriptive action verbs (DAV, e.g. A calls B) would represent the most concrete level of the model, followed by interpretative action verbs (IAV, e.g. A helps B), state verbs (SV, e.g. A likes B) and finally, adjectives (ADJ, e.g. A is friendly) for the most abstract level. Maass et al. (1989) have used measures of language abstraction in the description of behaviors between sports teams but also between religious groups. Following these authors, LIB is an implicit path to provide information about people's stereotypical beliefs without consciousness. Thus, people may be unable to exert intentional control over predicate abstraction (Franco & Maass, 1996 ; Maass et al. 1995 ; Maass et al., 1996 ; Maass, 1999).

Language abstraction has been studied in several ways as outcome: free completion, multiple-choice procedures, Likert-type scale (agreement on statement), content analysis (Douglas

et al., 2008; Li & Hills, 2021 ; Maass, 1999 ; Maass et al.,1989 ; Menegatti & Rubini, 2013; Scaffidi et al.,2020 ; Snefjella & Kuperman, 2015). Von Hippel et al., (1997) found that implicit and explicit measures of prejudice were uncorrelated. Following the authors, these measures are consistent with the idea that the degree of language abstraction used captures a valence of attitude that is not expressed by self-reported measures. Indeed, they found that implicit gender prejudice based on the LIB was correlated with implicit prejudice measure based on biased attributional processing but not correlated with explicit-prejudice measure (study 3).

Scaffidi et al. (2020) propose an experimental protocol in which participants were required to complete dialogues initiated between a Maghrebian and an Italian, either by taking on the role of the Italian or that of the Maghrebian or without additional information on the characteristics of the person whose dialogue is to be completed. The results indicate that completing dialogues from the perspective of a Maghrebian character (out-group) rather than an Italian one (in-group) led participants (Italians) to use less abstract terms in their dialogue than if they had not received this instruction. Thus, seeing from the perspective of an out-group member would result in expressing fewer prejudices through a lower degree of language abstraction. In the control condition (without information), the expression of prejudice through language abstraction aligned with typical findings in the literature (Maass, 1999; Maass et al. 1989; Wigboldus & Douglas, 2007). Indeed, participants employed more abstract terms to describe the negative behavior of the Maghrebian character, referring to stable and predictable predispositions, and less abstract terms to describe positive behavior, referring to more external or situational circumstances.

The aim of this research is to explore links between 1/implicit attitude measured by an SC-IAT, 2/ language markers: the degrees of language abstraction and 3/self-reported measure of racial prejudice. That's it in order to study the convergent validity between SC-IAT-P and the degree of abstraction of language based on LIB. In order to achieve this aim, some factors must be considered in the study of LIBs. Firstly, Schüller (2016) highlights that the 09/11 attacks in the United States

led to significant changes in the attitudes of the native German population, resulting in increased negative prejudice towards immigration. This effect is significant specifically among individuals with a low level of education, while those with a high level of education did not alter their attitudes. Studies showed that people with high levels of education would be more favorable to immigrants than people with low levels of education (Daniels & Von der Ruhr, 2003; Matthes & Marquart, 2015; Matthes & Schmuck, 2015). These results revealed once again the role of social desirability among participants with high levels of education when using explicit measures of attitude. Thus, when attitudes are indirectly measured, these effects may be reversed for highly educated individuals. Following Matthes and Schmuck (2015), these attitudinal differences between groups could be explained by the fact that people with higher levels of education would be more motivated to control their biases. Thus, it is necessary to consider education level as a control variable. However, social desirability, which may be responsible for a more pronounced normative bias among highly educated individuals, can be considered a potential moderator between implicit (i.e., LIB, SC-IAT-P) and explicit measures. Therefore, the more individuals are sensitive to social desirability and, consequently, to norms penalizing the expression of harmful attitudes, the less implicit and explicit measures would be correlated. Following Olsson (2023), when social desirability is prominent among individuals or within the given context, the association between implicitly measured negative racial biases (IAT) and support for the right-wing populist movement (RPP) becomes more pronounced for individuals exhibiting explicitly measured positive racial biases. An implicit assessment of racial biases consistently emerges as a crucial indicator of endorsement for the right-wing populist party (Maier et al., 2023; Olsson, 2023).

Firstly, to study the link between an SC-IAT-P and LIB we have to consider social desirability as moderator of the link between both implicit measures (LIB and SC-IAT-P) and explicit measure of attitude towards Maghrebians. Inconsistencies between the measurement of implicit and explicit attitudes may indicate sensitivity to social desirability, particularly when

assessing the link between anti-immigrant attitudes deemed undesirable and voting for a right-wing populist party (RPP). The implicitly measured right-wing populist political orientation (IAT) may be predictive of voting behavior. Furthermore, this effect persists even when controlling for traditional direct attitude measures (Bos et al., 2018). Secondly, political orientation is a moderator of the link between implicit and explicit attitudes especially when assessing attitudes towards a stigmatized ingroup. Indeed, it has been demonstrated that the political orientation of a message source influences, on one hand, the level of abstraction of the message and its evaluation by a recipient based on the congruence of their political orientation with that of the source. On the other hand, it also affects the degree of abstraction of the message depending on the subject matter. Meneghetti and Rubini (2013) manipulate the effect of congruence between the political orientation of the source and the target on the production of a more abstract or concrete discourse. The authors show that politicians generate a more abstract discourse when they believe they are addressing individuals with the same political orientation and a more concrete discourse when addressing individuals with a divergent orientation (Study 1). Additionally, activists perceive politically abstract speeches as more persuasive than concrete ones when they share the same orientations, and vice versa (Study 3). Burguet (2022) focuses on the role of political orientation in measuring explicit and implicit attitudes (via LIB) towards migrants among French participants. Results reveal that while left-wing participants have more favorable explicit attitudes toward migrants than right-wing participants, this difference is less pronounced at the implicit level. The analyses show a non-correlation between explicit and implicit measures on practically all positive and negative exogroup behaviors rated by left-wing people, while the opposite is observed for right-wing people. This results by indicating that individuals who are politically right-wing, conservative, and nationalistic would have more negative attitudes than those who are politically left-wing and would be subject to less normative pressure to express them.

Indeed, people who are politically right-wing, conservative and nationalistic would have more negative attitudes than people who are politically left-wing, and would be subject to less normative pressure to express them (Adorno, 1950; Berezin, 2009; Bierbrauer & Klinger, 2002; Esses & al, 2001, 2005; Feldman, 1988; , Hamberger & Hewstone, 1997; Kessler & Freeman, 2005; Malchow-Moller et al., 2008; Maier et al, 2023; Mayda, 2006; Miguet, 2008; Stephan et al., 1999; Voci & Hewstone, 2003). To study the link between an SC-IAT-P and LIB we have to consider political orientation as moderator of the link between both implicit measures (LIB and SC-IAT-P) and explicit measure of attitude towards Maghrebians.

The current research

Previous research led to conducting two studies to assess to what extent LIB can be an implicit measure of attitude towards sensitive objects as well as an implicit association test (SC-IAT-P). Self-reported measures can be biased by social desirability, especially when assessing prejudiced or socially sensitive attitudes such as political opinions or attitudes towards foreigners, in contrast to implicit measurements (Campbell & Stanley, 1963, Gawronski et al., 2015; Greenwald & Banaji, 1995). To highlight the relation between these two implicit measures and an explicit measure of attitude, we chose Maghrebians as stigmatized social group especially because French people demonstrate social desirability towards this group (Beauchemin et al., 2020; Simon, 2010; Lapeyronnie, 2008).

Before stating the hypotheses concerning the two studies in this research, some terminological and methodological clarifications are in order. The level of language abstraction (via the LIB)¹ is measured on the basis of the Linguistic Category Model (LCM: Semin & Fiedler, 1988; 1991) by an abstraction score, which responds to the outgroup hypothesis stipulating that people communicate more abstractly the undesirable than desirable behaviors of an outgroup (Maass et al.,

¹ From a terminological standpoint, although both studies measure the degree of prejudice through the levels of abstraction and concreteness in language, these levels theoretically pertain to the measurement of language abstraction via the LIB and are named as such in both studies.

1989, 1995). And it is measured on a proportion of more or less abstract terms, considering concrete terms as the least abstract. But if we consider that "*a positive behavior displayed by an in-group member will be described in relatively abstract terms, whereas the same behavior shown by an out-group member will be described in relatively concrete terms*" (Maass, 1999, p.80), It is therefore appropriate to measure the degree of prejudice underlying this measure of language abstraction (via the LIB) by two measures of level abstraction and language concreteness. However, this research addresses the hypothesis that the indirect measure via LIB should be as sensitive to negative exo group attitudes as the indirect measure via SC-IAT-P and both more sensitive than the direct measure. Thus, participants should display more prejudice towards North Africans via LIB when they place more importance on exogroup behaviors expressed negatively and more abstractly than concretely (LIB- score) and exogroup behaviors expressed positively and more concretely than abstractly (LIB+ score). For Study 1, LIB- and LIB+ scores will be calculated based on the higher proportion of choosing items more abstract than concrete for negative behaviors (abstraction measure) and the higher proportion of choosing items more concrete than abstract for positive behaviors (concreteness measure). In Study 2, LIB- and LIB+ scores will be calculated according to the higher rating of abstract than concrete items for negative behaviors (abstraction measure) and the higher rating of concrete than abstract items for positive behaviors (concreteness measure). As the level of abstraction related to negative behaviors (LIB-) and the level of concreteness related to positive behaviors (LIB+) measure the same concept (prejudice), we do not formulate a differentiated hypothesis. However, we assume a correlation between them and with the SC-IAT-P.

The aim of the study 1 was to assess the relationship between two indirect measures of attitude, one based on automatic memory associations (SC-IAT-P) and the other on language cues, specially the level of language abstraction (LIB). We not only hypothesize that these two measures are positively correlated but also that the SC-IAT is predictive of LIB, as it reflects a lower-level process than language abstraction. An SC-IAT-P (pen & paper version SC-IAT-P, Bardin et al.,

2014) and a measure of the language abstraction (LIB) inspired by Maass et al. (1989; 1995) were used as indirect measurement to assess attitude and prejudice towards Maghrebians. In addition, they filled surveys on political orientation, educational level, gender and age.

Furthermore, self-reported measures of racial prejudices are weakly correlated with the results of implicit association tests (see meta-analysis; Hofmann et al., 2005) as well as with the level of language abstraction (Von Hippel et al., 1997; Maass et al., 1995). Social desirability, on one hand, and political orientation, on the other hand, may act as moderators in the relationship between these indirect measures (linguistic biases and implicit associations) and self-reported measures (LIB: Burguet, 2022; Huber & Lapinski, 2006, IAT: Bardin et al., 2014, Schimmack, 2019).

Therefore, in study 2, we assess that social desirability on the one hand and political orientation affiliation on the other hand moderate the relationship between implicit and explicit attitudes towards Maghrebians (hypothèse 1). We also assess whether the interaction of social desirability and political orientation moderate the relationship between implicit and explicit attitudes towards Maghrebians (hypothèse 2). Implicit associations were measured with a SC-IAT-P informatized and the language of abstraction (LIB) by a procedure inspired from Von Hippel et al. (1997). Each hypothesis will be tested in different models. In one model the implicit attitude will be evaluated with SC-IAT-P, in the other model with LIB as antecedents. The preliminary results of the study 2, led us to conduct additional analyses beyond those required to test our original hypotheses. Specifically, the findings from the correlational analysis of study 2 lead to testing a moderated mediation model. Thus the path analysis is SC-IAT-P -> LIB -> Explicit attitude, controlling effects of political orientation and social desirability.

The two studies in this research have each been pre-registered and can be accessed at the following address: [https://aspredicted.org/X2W_VR7 (Study 1), https://aspredicted.org/JLC_DZN (study 2)]. Before participating in the research, participants were provided with an information

sheet and an informed consent form. Participants had the option to withdraw from the study at any time without any consequences. They were required to confirm their agreement for the retention of their data upon completing their participation. All data collected is anonymized.

This research has obtained a positive ethical approval from the research ethics committee of the Federal University of Toulouse (France) on number n°2020-29.

Study 1

Hypothesis

We hypothesize there is a correlation between scores of the two measures of implicit attitude towards Maghrebians : the SC-IAT-P and the level of abstraction used in the language (LIB). More precisely, we assume that SC-IAT-P is predictive of levels of abstraction after controlling political orientation, gender, age, education level and phototype of participants.

Method

Participants

Seventy participants took part in this study and were adults returning to study psychology (Masters level) and have completed the questionnaire during a social psychology course. All participants were thanked by the researcher, who answered all their questions. No incentive was offered. Among the 70 questionnaires collected, 9 were excluded from analyses because they have at least one of their two parents is of North African or Middle Eastern origin.

In the end, the analysis focuses on the surveys collected from 61 participants (12 men and 49 women, with a mean age of 37.46 [9.55]).

Materials

Level of language abstraction.

The language bias questionnaire (LIB) is a self-reported measure. The Linguistic Category Model (Semin & Fiedler, 1988) categorizes interpersonal predicates into four categories based on their level of linguistic abstraction. Previous researches (e.g., Assilaméhou-Kunz et al., 2020; Douglas & Sutton, 2010; Geschke et al., 2007; Wigboldus et al., 2000) relied on the manipulation of linguistic abstraction solely on the distinction between action verbs and adjectives. For ecological reasons, in this present series of studies interpretive action verbs as an intermediate position were chosen, describing a general class of behaviors with positive or negative connotations, rather than state verbs that refer to enduring states without a clear beginning or ending. This choice is particularly driven by the fact that the former is closer to the most concrete terms than the most abstract ones, which favors fewer abstract items. Then, the measure was composed of 24 items referring to 4 themes (employment, culture, economy, security). In each theme there were 3 positive and 3 negative items in 8 questions. For each question participants were asked to choose among 3 items "the one with which they most agree or least disagree". The questions were "Concerning employment, you would say:", "Concerning culture, you would say:", "Concerning the economy, you would say:", "Concerning the security, you would say:" The levels of abstraction have been reduced to 3 categories as mentioned above : 1/ Most abstract level: adjectives (ADJ), 2/Medium abstract level Interpretative verbs (IAV), 3/ Most concrete level : action verbs (DAV). Additionally, we have adopted the 4 themes (employment, culture, economy, and security) used in Burguet's study (2022) to draft our items. Each question included 3 items with three different levels of abstraction. For each question participant could check the answer "do not wish to answer". To measure the level of prejudice expressed in language, abstraction and concreteness scores are calculated, the former from negative items and the latter from positive items (i.e., LIB theory, Maass, 1999). In order to calculate an abstraction score and in accordance with Coenen et al., (2006), the negative items are weighted as follows: with descriptive action verb (1), with

interpretative action verb (2) and with adjectives (4). Concerning the concreteness score, positive items are weighted as follows: with descriptive action verb (4), with interpretative action verb (3) and with adjectives (1)². The score obtained is a prejudice score (PS). The higher the prejudice score for positive and negative items, the more negative the attitude towards Maghrebians people.

Pretest : *Abstraction language materiel*

As previously stated, the degree of prejudice was assessed by the level of language abstraction was assessed through statements describing either stereotyped (negative behaviors) or non-stereotyped (positive behaviors) according to 3 degrees of abstraction (DAV, IAV, ADJ). Pretest was conducted to verify that negative items measuring abstraction degrees are considered by the participants as being negative behaviors and positive ones as describing positive behaviors. Fifty-four students in second year communication (6 men and 48 women, with a mean age of 18.82 [0.91]) filled in a survey before a social psychology class. They were asked to indicate the extent to which 24 propositions can be considered negative or positive towards Maghrebini people on a scale from 0 (very negative) to 10 (very positive). All participants were thanked by the researcher, who answered all their questions. No incentive was offered. The item scores were all polarized and differ from the theoretical mean of the scale ($ps < .05$) except for 4. As there was no consensus on their polarity, they were modified to be more polarized than they were.

Implicit Association Test - P&P SC-IAT-P

The other measurement of implicit attitude was carried out using a paper-and-pencil implicit association test (p&p SC-IAT-P, Bardin et al., 2016). In this task participants must categorize as many words related to a target concept ("Maghrebians" represented by Maghrebians first names)

² "State verbs" are not measured, but are theoretically included between IAV and ADJ, and therefore should have had weight (3) for negative items and weight (2) for positive items.

with evaluative attributes (positive or negative words) in two opposing categories (I like *versus* I don't like) within 20 seconds. This test includes 2 steps, in one of them the target concept and positive words must be categorized in the "I like" category while negative words must be categorized in the "I don't like" one. In the other step, the target concept and negative words must be categorized as "I don't like" and positive words as "I like". For each step, there is a training phase. The steps are counterbalanced among subjects. The scores were calculated in accordance with Bardin et al. (2016)'s recommendations. A positive score corresponds to a positive attitude while a negative score corresponds to a negative attitude.

Political orientation. A slider formalized by a line of 10.5 cm at the left end was labeled "left" and at the right end "right". The instruction was to position a cross representing the political orientation on this line. The value of the political orientation was determined by measuring the distance between the beginning of the line and the cross inscribed by the participant. Thus, the higher the value, the more right-wing the participant's political orientation is..

Phototype. Fitzpatrick skin type assessment using the FCS (Fitzpatrick, 1988) was used.

Participants were asked to indicate their phototype by matching their skin color to the 6 proposed shades from lightest to darkest.

Control variables. Age, gender, parent's origin and education level was asked.

Procedure

Participants first completed the measurement of level of language abstraction (LIB), and subsequently implicit attitude towards Maghrebian (SC-IAT-P pen&paper). Finally, they indicated

political orientation, geographical origin of their parents, their phototype, level of education, gender and age.

Results

Preliminary analyses

Few of them have fully responded to the questionnaires related to the degree of language abstraction (prejudice score). Thus, only 28 participants checked at least one of the items from each of the 8 questions. For the remaining participants, at least one item out of 33 participants preferred to check "Do not wish to answer". Consequently, the two scores of language abstraction were therefore calculated for each participant by taking the average of the items actually responded. Furthermore, 1 participant did not respond to any positive items, and 7 others did not respond to any negative items. Therefore, the average scores could not be calculated for these participants. Three outliers were excluded from analyses in links with the SC-IAT-P score. The prejudice score is higher for positive items ($M=2.92$, $SD=0.60$) than negative items ($M=1.75$, $SD=0.50$), $t(50)=10.47$, $p<.001$, $d=0.79$.

Participants have a positive score of SC-IAT-P towards Maghrebians, $M=0.19$, $SD=0.35$, $t(58)=4.28$, $p<.001$, $d=0.26$.

All descriptive statistics and correlations are presented on table 1.

Table 1

Descriptive statistics and correlations between variables

	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7
1. SC-IAT-P (p&p)	0.14	0.26	-	-.05	-.04	-.08	-.03	.25	-.23
2. LIB +	2.97	0.61	-	-	-.02	.04	-.07	-.14	-.03
3. LIB -	1.81	0.57	-	-	-	-.13	-.11	.19	-.20

4. Age	37.46	9.55	-	-	-	-	-	-.15	-.11	.47**
5. Political orientation	3.36	2.59	-	-	-	-	-	-	.013	-.96
6. Phototype	3.11	0.99	-	-	-	-	-	-	-	-.27*
7. Educational level	7.14	1.13	-	-	-	-	-	-	-	-

Note. * $<.05$ (one-tailed), ** $<.05$ (two-tailed).SC-IAT-P= association test, LIB+ = prejudice score for the positive items, LIB- = prejudice score for the negative items.

No relation between LIB (- or +) was found with age, gender level of education, phototype and political orientation. The same pattern was found for SC-IAT-P. Contrary to our expectations, the hypothesis of a correlation between SC-IAT-P and LIB (positive or negative) was not confirmed. SC-IAT-P is not predictive of LIB- ($B=-.15$, $F(1,40)=0.78$, $p=.39$) and LIB + ($B=.06$, $F(1,45)=0.15$, $p=.70$) after controlling for political orientation, gender, age, education level and phototype of participants.

Discussion

The aim of study 1 was to demonstrate that the two implicit measures of attitude towards Maghrebians were not only correlated but also that the one based on automatic processes (SC-IAT-P) predicted the one based on the level of abstraction taking into account the possible effect of control variables .

Contrary to the hypothesis, the scores of the implicit attitude measure based on the evaluation of the strength of the association in memory and relying on automatic cognitive processes (SC-IAT) and those of the levels of abstraction measure (LIB) are not correlated. No correlation is found between these two variables and the participants' age, gender. The hypothesis of the predictivity of implicit attitude measured by SC-IAT-P on the degree of language abstraction is

also not validated. Several methodological limitations may explain the lack of a result regarding the relationship between SC-IAT-P scores and the degree of abstraction of language.

The measurement of Linguistic Intergroup Bias (LIB) was adapted following the usual methodological approach for this measure (Maass et al., 1995; , 1989; Semin & Fiedler, 1989). In these experiments, participants viewed drawings depicting members of different groups exhibiting either positive or negative stereotyped behaviors. Participants then had to select one of the four provided statements that best described the protagonist's behavior. These descriptions represented the four levels of abstraction from Semin and Fiedler's Linguistic Category Model (1989), ranging from the most concrete (descriptive action verbs) to the most abstract (adjectives). However, it is possible that the way of measuring implicit attitude by exclusively choosing one description with a level of abstraction (LIB) may not fully capture the precise degree of the measured attitude on the one hand and exclude the potential agreement of the participant with other descriptions on the other hand. To address this, it is relevant to provide an evaluation of each proposed description.

Moreover, implicit measures are supposed to vary less according to the environment in which they are elicited than explicit measures of attitude, especially when it comes to attitudes towards sensitive objects influenced by social desirability, such as Maghrebians in France. Thus, if participants are concerned about social desirability and/or if the environment “encourages” adhering to normative attitudes, a weak correlation is usually found between implicit and explicit measures (Greenwald et al., 2009). Consequently, implicit measures should reflect an individual’s personal attitude whereas self-reported measures are potentially influenced by social norms and social context. However, it is not possible to definitively rule out a potential effect of social desirability on the SC-IAT-P scores. So, the implicit association test (SC-IAT-P) was conducted using its pen-and-paper version. This has several limitations compared to its computerized version. Among these limitations, the most significant one is it might lack the precise timing capabilities of the computerized version. Reaction times are critical for assessing implicit biases and the pen-and-

paper version records this time with lower precision. Thus, participants might not have stopped the task when the experimenter requested, leading them to take more time to respond. Taking more time to respond than instructed could result in participants' answers relying more on controlled cognitive processes than automatic processes. An effect of social desirability cannot be ruled out on the scores of SC-IAT-P.

Concerning LIB measurement, as soon as the survey was distributed, a large number of participants contested the measure of abstraction of language for the negative items and expressed their discomfort at having to make choices exclusive of each other. A significant number did not answer the 4 questions about negative items. Hence, 54% of the participants chose not to respond to at least one of the options in the LIB descriptions. The collective manifestation of rejecting negative items and dissatisfaction could lead participants to carefully consider their evaluations of the LIB items and exhibit social desirability bias in their responses.

Thus, before rejecting the hypothesis that the implicit measure using the SC-IAT-P would predict the implicit measure related to levels of language abstraction, we propose in a second study to address the methodological and contextual issues that have been raised. In accordance with the methodology of Von Hippel et al., (1997), instead of making a forced choice among three descriptions with varying levels of abstraction, we will ask participants to evaluate each of the proposed descriptions on Likert-type scales. These measures will not only introduce more variability in responses for each item but also avoid forcing participants to reject items to retain only one. Thus, we will have a relative measure of participants' agreement degree on each of the proposed items, based on the level of abstraction of each item.

Control variables from Study 1 will be retained in Study 2 except for the political orientation. Indeed, Burguet (2022) focuses on explicit and implicit attitudes towards migrants among French participants. Results reveal that while left-wing participants have more favorable

explicit attitudes toward migrants than right-wing participants, this difference is less pronounced at the implicit level. The analysis indicates that attitudes are not correlated among left-wing participants in contrast to right-wing participants. Consequently, according to Burguet (2022) this variable will be considered as a moderator of the links between implicit (SC-IAT-P and LIB) and explicit attitudes in part of analysis of study 2 and not systematically as control variable like in study 1.

Additionally, it is relevant to include an explicit measure of prejudice and assess a potential effect of social desirability on the relationship between implicit and explicit measures. This explicit measure of prejudice towards Maghrebians will also allow for comparing effect sizes of different antecedents on the explicit measure.

Study 2

The aim of this second study was to correct some methodological and theoretical issues encountered in Study 1. It aimed to highlight the relationship between implicit attitude and degree of linguistic abstraction (via LIB) while considering the level of social desirability using measures that provide a greater variability of responses than in Study 1. In Study 1, the measure of degree of language abstraction consisted of selecting an item from among others, which did not allow participants to qualify their degree of adherence to the item's formulation. This led a significant number of them to refrain from answering. Von Hippel et al. (1997) propose a measure of prejudice based on LIB (e.g., Maass et al., 1995) to quote items on a scale. For example, participants have to evaluate four statements (describing an article that they have read) corresponding to four levels of abstraction. They indicated the extent statements describing the article on 10-point Likert-type. This type of measure would allow for more variability in responses and would prevent participants from refusing to answer because they could qualify their agreement with the different statements.

Furthermore, the measurement of implicit association will be conducted using the computerized version (SC-IAT-P, Bardin et al., 2014) to address the limitations of study 1 concerning the pen and paper version.

Hypothesis 1

We assumed that social desirability on the one hand and political orientation on the other hand moderate the relationship between implicit and explicit attitudes towards Maghrebians after controlling for gender, age, education level and phototype.

Hypothesis 2

We assumed that the interaction of social desirability and political orientation moderate the relationship between implicit and explicit attitude towards Maghrebians after controlling for gender, age, education level and phototype of adult participants.

For all hypotheses, implicit attitudes will be measured with 1/ an implicit association test (SC-IAT-P), 2/ the level of language abstraction (LIB). Each hypothesis will be tested in two different models. In one model the implicit attitude will be evaluated with SC-IAT-P, in the other model with LIB.

Method

Participants

One hundred and thirty-two adults took part in the research and were recruited in four different ways. Some of them were adults returning to study psychology (Masters level) and participated one by one by videoconference due to COVID crisis. The visio conference allowed the experimenter to give the instructions and then the participant carried out the task. At the end, the experimenter debriefed them via visio conference. Other participants were third year communication-students and answered the questionnaire before a social psychology class. All participants were thanked by the researcher, who answered all their questions. No incentive was

offered. Among the 132 questionnaires collected, 24 were excluded from the analysis because they did at least one of these characteristics: not complete the SCIAT-P (implicit test) due to technical problems (n=8), have at least one of their two parents is of North African or Middle Eastern origin (n=15) or refuse to allow his data to be stored (n=1). At the end, the analysis focuses on the questionnaires collected from 118 participants (21 men and 97 women, with a mean age of 28.88 [11.44]). Furthermore, 15 participants did not provide information about their educational level.

Materials

Explicit attitude. Explicit attitude towards Maghrebians will be measured with the Blatant and Subtle Prejudice Scale (BSPS; Pettigrew & Meertens, 1995). It is a self-reported measure with a 10-point Likert-type scale. It consists of 20 statements which the participants rate on a scale from 0 to 9. The reliability analysis revealed good internal consistency ($\alpha=.89$), which allowed an attitude to be computed: The higher the score, the more positive the attitude towards Maghrebian is.

Social desirability. The BIDR (Balanced Inventory of Desirable Responding; Paulhus, 1998) is the most widely used tool to assess social desirability. We will use a derived form translated in French: DS 36 (Tournois et al., 2000). It is a self-reported measure with a 7-point likert-type scale. It consists of 36 statements which participants rate on a scale from 1 ("Strongly disagree») to 7 scale («strongly agree"). The reliability analysis revealed good internal consistency ($\alpha=.86$), which allowed a social desirability score to be computed: The higher the score, the more participants are likely to demonstrate social desirability.

Political orientation. Political orientation was measured by using a cursor ranging from "left" to "right". To the left of the cursor the value will be 0 and 100 to the right. The incrementation of the cursor was 1.

Implicit association test (SC-IAT-P). The SC-IAT-P (Bardin et al. 2014) is a variant of the IAT (Implicit association test, Greenwald et al., 1996). It is a computer-based categorization task which is conducted in two measurement phases based on combined categorizations. In the first phase, the participant is asked to categorize the target concept and the positive evaluative attributes on one key and the negative evaluative attributes on another key (e.g. I like + Maghrebians *versus* I dislike). In the second, the association is reversed, with the participant having to associate the target concept and the negative evaluative attributes on one key and categorize the positive evaluative attributes on the other key (e.g., I dislike + Maghrebians *versus* I like). The items of the target concept "Maghrebians" were eight first names of Maghrebians origin that had to be categorized in the target concept "Maghrebians". Feedback was given to the participant for each answer. The feedback was a red cross in case of an error and a green circle in case of a correct answer. This feedback appeared in the center of the screen before moving on to the next item. The items to be categorized as "I like" were: holiday, victory, gift, happiness, festive, flower, offer, happy, beauty and sweet. The words to be categorized as "I don't like" were: war, pain, accident, infection, funeral, vomit, purulent, sad, torture and hideous. To perform the categorization, participants had to use the 'a' and 'p' keys on the keyboard. Inquisit 6.00 web software was used to collect the data. The result of the test is a score. If the response time is faster for the association Maghrebians/I like than for the association Maghrebians / I dislike, the implicit attitude is positive. If the association of Maghrebians /I like is the fastest, the implicit attitude is negative. Thus, the higher the SC-IAT-P score, the more positive the attitude towards Maghrebians.

Level of language abstraction. The measure of level of language abstraction (LIB) is a self-reported measure with an 11-point Likert-type scale. Participants rate 24 statements validated in studies 1 and 2 on a scale from 0 to 10 scale of "Strongly disagree" to "strongly agree". The 24 statements are grouped into 4 themes: employment, culture, economy and security. There are three

levels of abstraction either positive or negative per theme: 1/ Most abstract level: adjectives (ADJ), 2/Medium abstract level : Interpretative verbs (IAV), 3/ Most concrete level : action verbs (DAV). In order to make the scores interpretable, the LIB positive items were coded in such a way as to invert the scale. Then, in accordance with Coenen, Hedebow and Sernin (2006) instructions, the scores for the positive items ADJ ($\alpha=.83$), IAV ($\alpha=.79$) and DAV ($\alpha=.76$) were weighted as follows ADJ*1, IAV*3 and DAV*4 respectively. The scores for the negative items ADJ ($\alpha=.82$), IAV ($\alpha=.76$) and DAV ($\alpha=.59$) were weighted as follows ADJ*4, IAV*2 and DAV*1 respectively. Thus, a prejudice score was obtained for the positive items (LIB+) and a second prejudice score for the negative items (LIB-).

The higher the prejudice score is, the more negative the attitude towards Maghrebin people.

Phototype. Fitzpatrick skin type assessment using the FCS (1988) was used. Participants were asked to indicate their phototype by matching their skin color to the 6 proposed shades from lightest to darkest.

Control variables. Age, gender, parent's origin and education level was asked.

Procedure

Participants first completed the measurement of attitude explicit towards Maghrebians (BSPS), and subsequently a score of racial prejudice *via* levels of language abstraction (LIB), social desirability, political orientation measures and SC-IAT-P. Finally, they indicated the geographical origin of their parents, their phototype, gender, educational level and age.

Results

Analytical procedure

The estimated direct, interaction and indirect effects for the mediation or moderate models were generated by a macro for SPSS, PROCESS (Hayes, 2018) which uses an OLS regression procedure. Conditional indirect effects are produced with a pick-point approach (+ \ - 1 SD from moderator's mean). Bootstrap confidence intervals (bias-corrected) are based on 10,000 bootstrap.

Preliminary analyses

Karpinski and Steinman (2006) recommend excluding data from participants with over 20% of errors for the computerized SC-IAT. Consequently, the data from the 10 participants who returned more than 20% of errors were excluded from analysis on this measure as well as one participant who had outlier data.

Preliminary analyses (simple regressions) showed an effect regarding age on social desirability indicating that the older the participants, the more inclined they were to show social desirability, ($B=.47$, $F(1,116)=32.31$, $p<.001$). An effect regarding age was found also on political orientation ($B=.23$, $F(1,116)=6.67$, $p=.011$) indicating the older the participants, the more right-wing they are. No effect was found regarding age on the other variables (SC-IAT-P, LIB, BSPTS) and regarding the gender and phototype of participants. Effect of level education was found only on social desirability ($B =.41$, $F(1,103)=20.68$, $p<.001$), the higher the educational level, the more inclined they were to show social desirability.

As a consequence of these preliminary analyses, analyses related to social desirability and political orientation are presented including age and educational level as covariates.

Participants have a positive explicit attitude towards Maghrebians ($M=6.80$, $SD=1.32$), $t(117)=14.75$, $p<.001$, $d=1.32$. The prejudice score is higher for positive items (LIB+, $M=7.97$, $SD=5.31$) than negative items (LIB-, $M=5.93$, $SD=4.27$), $t(117)=-5.43$, $p<.001$, $d=4.08$. Participants have a positive score of SC-IAT-P towards Maghrebians, $M=0.07$, $SD=0.35$, $t(102)=1.98$, $p=.05$, $d=0.35$.

Preliminary analysis show an effect of implicit attitude measured by the SC-IAT-P on explicit attitude (BSPS, $B=.36$, $F(1,101)=14.76$, $p<.001$), on (LIB+) ($B=-.32$, $F(1,101)=11.78$, $p=.001$) and (LIB-) ($B=-.25$, $F(1,101)=6.50$, $p=.012$). The higher their score SC-IAT-P, the lower their score for LIB+ and LIB-. Preliminary analyses show also an effect of (LIB+), $B=-.68$, $F(1,116)=101.53$, $p<.001$ and LIB-, $B=-.83$, $F(1,116)=260.57$, $p<.001$ on BSPS. The higher their score for LIB+ and LIB-, the lower their BSPS scores.

Descriptive statistics and correlations are presented on table 2.

Table 2

Descriptive statistics and correlations between variables

	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9
1. SC-IAT-P	0.06	0.37	-	.36**	-.32**	-.25*	.10	.11	-.12	.09	.07
2. BSPS	6.77	1.35		-	-.68**	-.83**	-.10	-.05	-.45**	-.02	-.03
3. LIB +	8.17	5.44			-	.66**	.11	.07	.42**	.03	.13
4. LIB -	6.18	4.40				-	.12	.09	.46**	.02	.08
5. DS 36	4.24	0.73					-	.47**	.12	.00	.40**
6. Age	28.51	10.89						-	.23*	.09	.70**
7. Political orientation	40.22	24.72							-	.05	.12
8. Phototype	3.23	1.07								-	.21*
9. Educational level	7.12	1.25									-

Note. SC-IAT-P= association test, BSPS = explicit attitude, LIB+ = abstraction score for the positive items, LIB-= abstraction score for the negative items, DS36= social desirability, * $p<.05$. ** $p<.01$

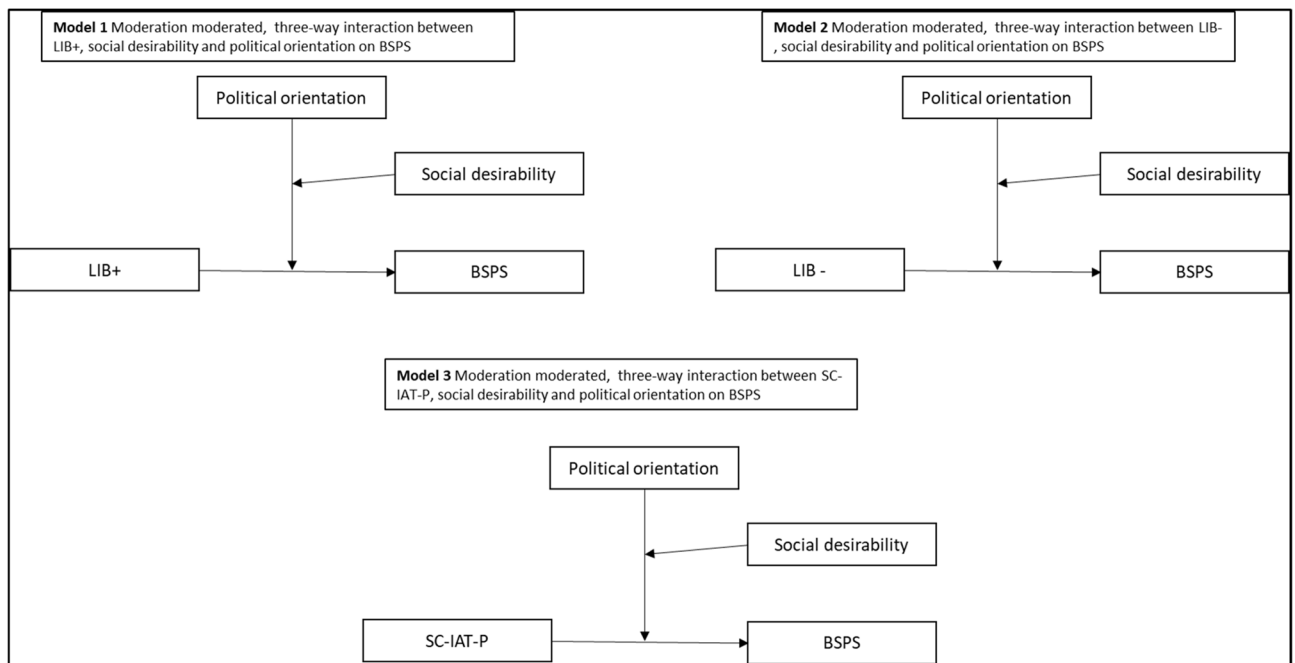
Main models and results

To assess hypothesis 1, six analysis were conducted, one for each effect of interactions on explicit attitude: H1a) SC-IAT-P*political orientation, H1b) SC-IAT-P*sociability, H1c) LIB+*social desirability, H1d) LIB-*social desirability, H1e) LIB+*political orientation, H1f) LIB-*political orientation.

To assess hypothesis 2, three models were tested (fig. 1)

Figure 1

Moderated Moderations models (Hypothesis 2)



Note. SC-IAT-P= implicit association test, BSPS = explicit attitude towards Maghrebians, LIB+ = abstraction score for the positive items, LIB- = abstraction score for the negative items

Simple moderation analyses.

The moderation analysis (H1a and H1b) revealed that the effect of implicit attitude (SC-IAT-P scores) on explicit attitude depends on political orientation (table 3, graph 1) and social desirability (table 4, graph 2)

Table 3

Interaction between SC-IAT-P and political orientation on BSPS

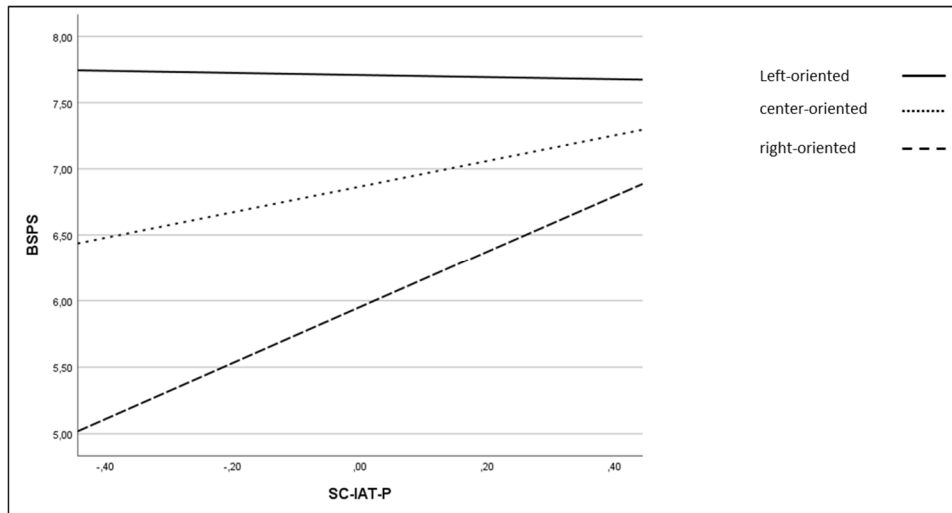
Consequent			
BSPS			
Antecedents	Coeff.	SE	p
SC-IAT-P	1.06	.32	.0014
Political orientation (PO)	-.03	.0047	<.0001
Interaction (SC-IAT-P*PO)	.041	.013	.0026
Age	.022	.015	.14
Educational level	-.07	.12	.57
constant	6.70	.65	<.0001
$R^2=.46$			
$F(5, 86)=14.82, p<.0001$			

Note. SC-IAT-P= implicit association test, BSPS = explicit attitude towards Maghrebians

Probing interaction with the pick-a-point approach (mean and ± 1 SD from mean) reveals conditional effects of implicit attitude for center-wing ($p=.0014$) and right-wing ($p<.0001$) political orientation (graph 1). Political orientation is a quasi-moderator because it is also a predictor (independent variable) of explicit attitude. This pattern of results is interpreted as an overcompensation between explicit and implicit measures in left-oriented people compared to central and right-oriented people. Thus left-oriented participants have a positive explicit positive attitude even if they have a negative implicit attitude.

Graph 1

Conditional effect of SC-IAT-P at values of political orientation



Note. SC-IAT-P= implicit association test, BPS = explicit attitude towards Maghrebians

Table 4

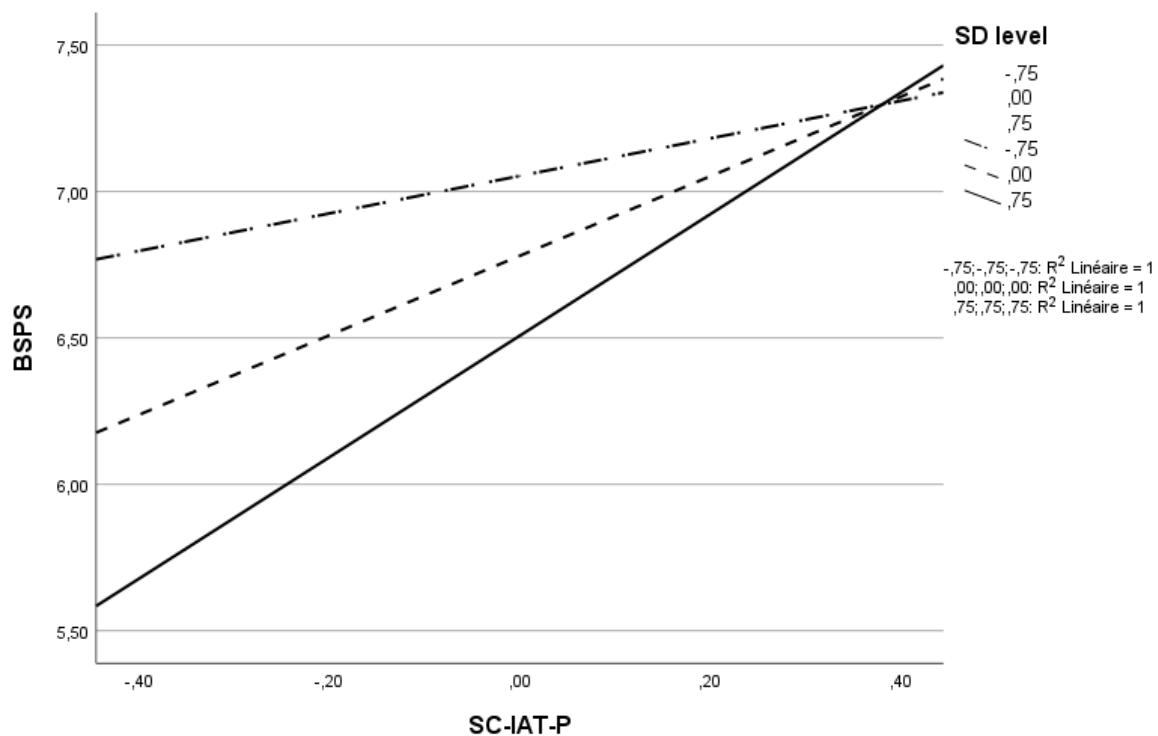
Interaction between SC-IAT-P and Social Desirability on BPS

Consequent			
BPS			
Antecedents	Coeff.	SE	p
SC-IAT-P	1.36	.40	.0009
Social Desirability (SD)	-.36	.21	.09
Interaction (SC-IAT-P*SD)	.96	.48	.048
Age	.022	.015	.14
Educational level	.023	.15	.88
constant	6.65	.86	<.0001
$R^2=.43$			
$F(5, 86)=3.81, p<.0037$			

Note. SC-IAT-P= implicit association test, BPS = explicit attitude towards Maghrebians, SB= social desirability

Graph 2

Conditional effect of SC-IAT-P at values of social desirability



Note. SC-IAT-P= implicit association test, BPS = explicit attitude towards Maghrebians, SD = social desirability

Probing interaction with the pick-a-point approach (mean and ± 1 SD from mean) reveals conditional effects of implicit attitude for medium ($p=.0009$) and high ($p<.0001$) levels of social desirability (graph 2). Direct effect of Social desirability on explicit attitude is not significant. Thus concerning participants who have a high or medium level of social desirability, the more they have a positive implicit attitude towards Maghrebians, the more they self-reported an explicitly positive attitude compared to the responses of participants with a low level of social desirability.

Concerning levels of language abstraction (H1c and H1d), no significant interaction was found between LIB + (or LIB-) and social desirability on BPS while there is direct effects of LIB (LIB+ : $B= -.17$, $SE=.02$, $p<.0001$, 95% CI [-0.20, -.13]; LIB- : $B=- .26$, $SE=.02$, $p<.0001$, 95% CI [-0.30,- .23.]) on BPS. Thus, the higher LIB+ (LIB-), the lower the BPS score is.

Same patterns of results was found concerning predictors effects (LIB+, LIB-) on BPS (H1e and H1f) when the moderator is political orientation (respectively, LIB+ : $B= -.13$, $SE=.02$,

$p < .0001$, 95% CI [-.17, -.09]; LIB- : $B = -.25$, $SE = .02$, $p < .0001$, 95% CI [-.30, -.21.] while no significant interaction is found neither LIB+ or LIB- with political orientation on BSPS score.

Concerning moderate moderation models (1, 2 and 3: hypothesis 2x) regressions coefficients for interactions were not statistically significant for models which have prejudice scores (LIB- / LIB+) as antecedents respectively; $b = -.0009$, $t(99) = -0.95$ (model 1), $p = .34$; and $b = -.0011$, $t(99) = -1.23$, $p = .22$.

The moderate moderation of model 3 (SC-IAT-P*PO*SD on BSPS) was significant, $\Delta r^2 = .026$, $F(1, 81) = 4.21$, $p = .042$ indicating 2.6% of additional variance explained by this three-way interaction. Conditional effects of the SC-IAT-P at values of the social desirability and political orientation conducted to conclude that there is an effect for participants who have a medium or right political orientation and medium or high level of social desirability (table 4, graph 3).

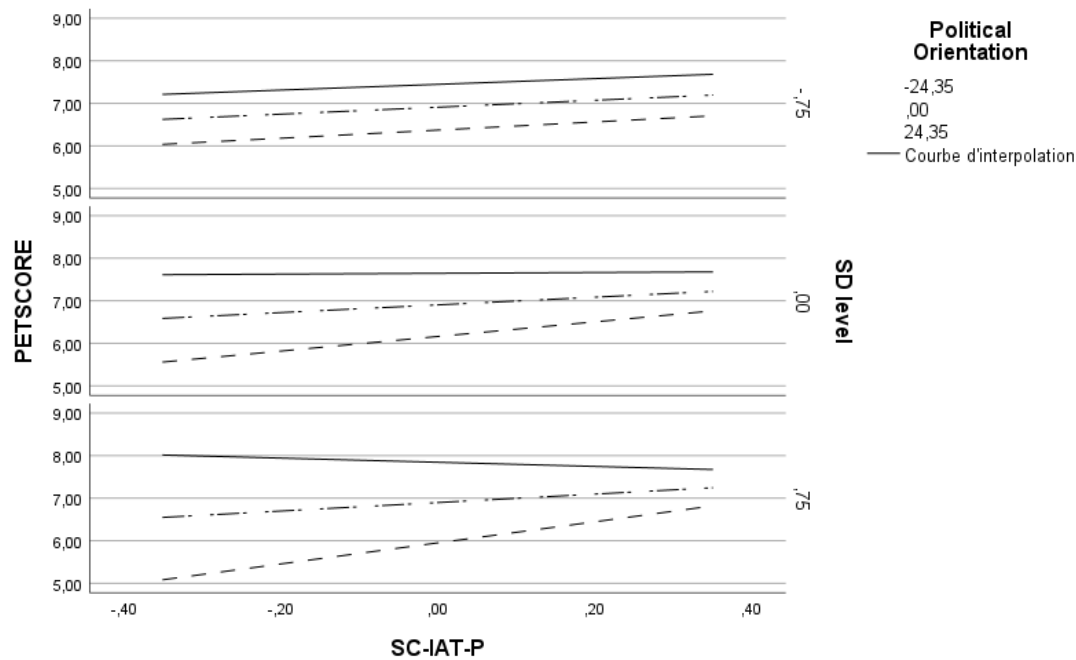
Table 5*Conditional effects of SC-IAT-P on BPSPS at values of political orientation and social desirability*

PO	SD level	Effect	SE	<i>t</i>	<i>p</i>	LLCI	ULCI
Left	Low	.67	.62	1.08	.28	-.56	1.91
	Medium	.93	.47	.20	.84	-.83	1.02
	High	-.49	.67	-.73	.47	-1.81	.84
Center	Low	.81	.45	1.79	.08	-.09	1.72
	Medium	.90	.33	2.77	.007**	.2542	1.55
	High	.99	.44	2.28	.03*	.13	1.86
Right	Low	.95	.72	1.33	1.19	-.47	2.38
	Medium	1.71	.47	3.67	.0004**	.78	2.64
	High	2.47	.56	4.43	<.0001**	1.36	3.58

Note. PO= political orientation, SD = social desirability, * $p < .05$. ** $p < .01$.

Graph 3

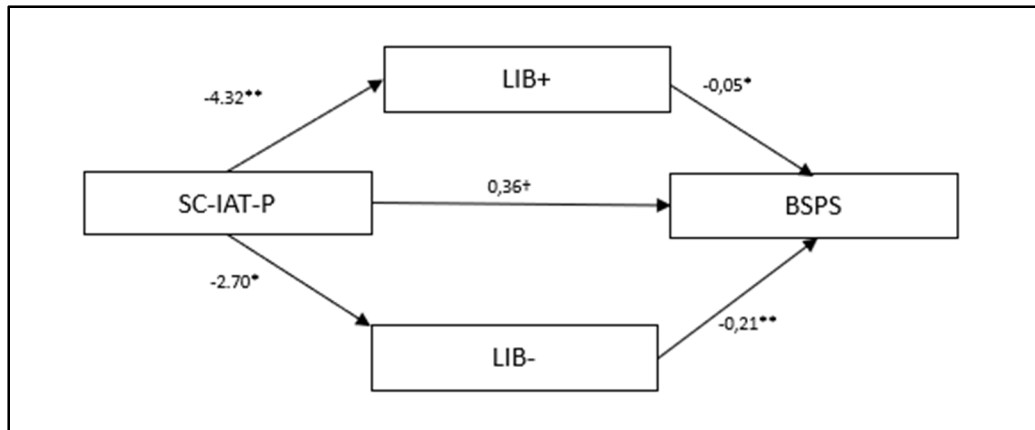
Conditional effects of the focal predictor at values of political orientation and social desirability



Because the three ways interaction of Model 3 was significant and no models concerning LIB (model 1 : LIB+*OP*SD on BSPS; model 2 : LIB-*OP*SD on BSPS) other exploratory analyses were conducted. Indeed, the relation between SC-IAT-P and BSPS seems to be more affected by political orientation and social desirability than LIB+/ LIB-. This pattern of results could stem from implicit attitude assessed by automatic association (SC-IAT-P) that is a lower level of cognitive process than language abstraction. This new hypothesis is also supported by the correlation levels between in one hand SC-IAT-P and BSPS (table 2, $r = .36$) and in the other hand LIB+/LIB- and BSPS (table 2, respectively: $r = -.68$ and $r = -.83$). Thus, language abstraction seems to be a measure closer to self-reported measure than SC-IAT-P. Consequently, we assume that LIBs (+ and -) are mediators of the relation between SC-IAT-P and BSPS (fig 4) when political orientation and social desirability are controlled as well as age and educational level.

Figure 2

Parallel moderated mediation of the relation between SC-IAT-P and BSPS by LIB- and LIB+ as mediators



Note. SC-IAT-P= implicit association test, BSPS = explicit attitude towards Maghrebians, LIB+ = abstraction score for the positive items, LIB-= abstraction score for the negative items

The total effect of SC-IAT-P on BSPS is significant, $t(86)=3.32$, $p=.0013$, 95%CI [0.45; 1.79] and partitioned on a direct effect and two simultaneous indirect effects by LIBs. Direct effect is no significant, $t(84)=1.60$, $p=.11$, 95%CI [- 0.07; 0.80] whereas the indirect effects are significant, respectively for LIB + and LIB - : $coeff.=0.20$, $SE=.10$, 95%CI [0.04; 0.42] ; $coeff.=0.57$, $SE=.25$, 95%CI [0.12; 1.11]. A pairwise comparison was conducted to assess an eventual difference between the two indirect effects. Thus, a specific indirect effect contrast ($C = LIB- minus LIB+$) was used (Hayes 2018) to determine whether the valence of items (positive *versus* negative) can have a differentiated effect on BSPS (PROCESS, Hayes, 2018). The indirect effects of SC-IAT-P on BSPS through LIB- (and LIB+) are not different from each other $C=0.37$, $SE=0.24$, 95%CI [-0.05;0.89].

Discussion

It was expected that the measures of implicit attitude would have a weak correlation with the explicit measure of attitude. This expectation is partially confirmed. Indeed, there is a moderate correlation between the scores of SC-IAT-P and BSPS ($r = .36$), while there is a strong correlation

between BSPS and LIB+ ($r = -.68$) and LIB- ($r = -.83$). Furthermore, it was also expected that the measures of implicit attitude would strongly correlate with each other since they are supposed to measure the same construct implicitly. Surprisingly, SC-IAT-P moderately correlates with LIB+ ($r = -.32$) and weakly with LIB- ($r = -.25$), while LIB+ and LIB- strongly correlate with each other ($r = .66$). These results support the interpretation that the different implicit measurements assess the same construct but in different ways.

As expected, the scores of prejudices related to positive and negative language descriptions (LIB) correlate with those of the political orientation measure. Contrary to expectations, the scores on the implicit measure related to automatic processes (SC-IAT-P) do not correlate with political orientation. No moderation by political orientation or social desirability (simple moderation model or moderated moderation model) was found on the link between the implicit attitude measure (LIB) and the explicit one (BSPS). However, political orientation on one hand and social desirability on other hand moderate the relationship between SC-IAT-P and BSPS. Assuming our assumptions we tested the effects of moderated moderation on the links between SC-IAT-P or LIB and BSPS for orientation political (M1) and social desirability (M2). The three-way interaction concerning LIB as antecedent was not significant while the three-way interaction concerning SC-IAT-P was. Thus, moderation of political orientation by social desirability influenced the relationship between the SC-IAT-P measure and the explicit attitude measure, but did not affect the relationship between LIB measures (LIB+ and LIB-) and the explicit measure. The analysis of this moderation (± 1 SD from the mean of the moderator) leads to the conclusion that participants at a low level (i.e., left-leaning) of this moderating variable express a positive explicit attitude towards Maghrebians, irrespective of the valence of their implicit attitude measured with SC-IAT-P, and this holds true regardless of their level of social desirability.

In contrast, participants with medium or high levels of social desirability and those with a center or right political orientation demonstrate greater alignment between their explicit and implicit attitudes toward Maghrebians. It appears that those politically left-leaning are less likely to explicitly express their prejudice, irrespective of the level of social desirability experienced, whereas individuals from the center or right politically adapt the expression of their explicit prejudice to their implicit prejudice, especially when exhibiting a high or medium degree of social desirability. In summary, it can be suggested that it is more challenging for left-leaning individuals to explicitly express their prejudice compared to those in the center and on the right.

A parallel mediation model was tested and revealed a prediction of SC-IAT-P on LIB+ and LIB- considered as mediators in the relationship between SC-IAT-P and BSPS. These results support the hypothesis that the implicit attitude measured by the SC-IAT-P may be based on lower levels of processing, while the language biases expressed through the degrees of abstraction would result from higher-level processing.

Discussion and conclusion

In accordance with our hypotheses, LIB+ and LIB- are strongly correlated. More specifically, in Study 2, LIB+ scores are more strongly correlated with SC-IAT-P scores than LIB- scores. This result is surprising because one might have thought that agreeing more with the outgroup's negatively stereotyped behaviors, expressed more abstractly than concretely, reflects greater prejudice than agreeing with its non-stereotyped, positive behaviors, expressed more concretely than abstractly. Therefore, it was expected that scores related to LIB- would show a stronger correlation with those of the SC-IAT-P and a weaker correlation with those of the explicit measure compared to scores related to LIB+. However, in both studies, participants showed greater hesitation in expressing opinions about negative elements than positive ones. It is highly likely that

the valence of negative elements was perceived as explicitly associated with negative prejudices and, consequently, more subject to social desirability concerns than positive elements.

Despite the difficulty in controlling for similar content with different valences in items (Douglas & Sutton, 2006), research studies (Assilaméhou et al., 2020; Douglas & Sutton, 2010) indicate that the LIB effect is independent of the valence of its items. These findings even suggest that responses to negative items may have influenced responses to positive items. One could indeed consider that the negative items may have made the participants more aware, and in order not to appear racist, they may have polarized their responses to the positive items to demonstrate a positive attitude towards Maghrebians people. Therefore, it can be suggested that the implicit measurement of attitude through LIB would be more appropriate when presented exclusively with positively oriented items, offering a subtler means to gauge racial biases.

In some previous research, beyond manipulating the degree of language abstraction, procedures involved participants viewing positive and negative behaviors (such as photos or videos) before making choices among items with different levels of linguistic abstraction (e.g., Wakslak et al., 2014) or evaluating them using attitude scales (e.g., Von Hippel et al., 1997).

This procedure highlights, in a more subtle and indirect manner, the connection between a person's identity (e.g., Maghrebians) and the expression of a prejudice towards them. Therefore, presenting only positive items in this type of procedure could enable participants with more discriminatory tendencies to express their disagreement or, at the very least, to contextualize the significance of these items with a more nuanced evaluation.

The results of Study 1 do not confirm the predictability of SC-IAT-P measures on LIB measures, whereas this is the case in Study 2. Methodological issues regarding the measurement tools were identified in Study 1, and these were appropriately addressed in Study 2. Furthermore, Study 2 demonstrated predictability of both LIB measures (LIB+ and LIB-) on the explicit measure.

A parallel mediation model was validated showing that the relationship between the implicit measurement of SC-IAT-P and the explicit measurement of BSPS was mediated by the implicit measurement of LIBs (LIB+ and LIB-) in parallel. Analyses indicated a total mediation. Thus, the direct effect was not significant while the two indirect effects were.

Indeed, it seems that the attitude measurement through SC-IAT-P is based on lower-level processing compared to the attitude measurement based on LIB, particularly with LIB-, which is even more pronounced than LIB+. The scores of LIB are more correlated with those of SC-IAT-P. It would be interesting to use LIB as an attitude measure in the analysis of language productions, such as political speeches or verbal responses in interviews. *“Although there are undoubtedly more powerful indirect measures of prejudice (such as those based on reaction time to primed stimuli-see Fazio, Jackson, Dunton, & Williams, 1995), LIB-based prejudice measures have the advantage that they can be used very easily in almost any setting without requiring laboratory space or computer equipment; more importantly, these procedures can be applied to subject populations”* (Maass, 1999, p.115). This is especially relevant if the topics discussed are sensitive and if the speakers identify themselves as politically left-leaning. Following Maass (1999), if the use of different levels of linguistic abstraction is often unconscious, it may be interesting to analyze them in normative contexts where the explicit expression of hostility is difficult to articulate or typically inhibited by certain populations. It is possible that in this research, the expression of hostility towards Maghrebians may have been even less explicitly articulable when individuals declared themselves politically left-leaning. This finding aligns with Burguet's (2022) study, which reported no correlation between implicit and explicit measures among politically left-leaning individuals, unlike those on the right. However, the interpretation proposed in that study suggested that left-leaning individuals were more susceptible to social desirability, a trend not supported by our results. Left-leaning individuals seem to consistently exhibit a politically correct attitude, irrespective of prevailing norms of the ingroup. It is possible that they may experience cognitive dissonance in

expressing explicit attitudes that go against their social identity as 'left-wing people' regardless of the context. In summary, social desirability has an impact on moderation by political orientation only for individuals with a medium or high level of social desirability among right-leaning or centrist individuals. In other words, the level of desirability among left-leaning individuals has no effect on the moderation by political orientation on the relationship between SC-IAT-P and BSPS. Regardless of the context, left-leaning individuals express a positive explicit attitude towards Maghrebians. Conversely, a more pronounced effect of social desirability is observed among right-leaning individuals. It may seem peculiar to conclude that right-leaning individuals with high social desirability have a stronger link between SC-IAT and BSPS, but we may currently be witnessing a shift in the normative window of prejudice (Mendiburo-Seguel & Ford, 2019) with unreserved right-leaning individuals expressing their opinions, especially when these opinions are considered socially desirable within the ingroup.

Limitations and Further Studies

There are methodological limitations in both studies. Those from Study 1 were addressed in Study 2, particularly regarding the paper-based nature of the measures and the response choices for LIB measurement. Therefore, it is advisable to employ a more precise and rigorous measure of attitudes through a computerized version of an implicit association test, as well as a more relevant measure of LIB through item measurement scales. The Study 2 is not without methodological limitations either. For instance, the SC-IAT-P is conducted under time constraints, while the LIB measures are not. It would be wise to introduce cognitive charge for the LIB measure as well. Thus, participants would have fewer cognitive resources available to control their responses. Considering the more problematic results regarding the negative items of LIB, which appear to be more explicit in expressing prejudice towards Maghrebians, it is advisable to focus on using positive items from LIB to obtain a subtler and less direct measure of racial prejudice.

Conclusion

This research confirms the value of LIB as an indirect measure of attitude. The results indicate that LIB serves as a more subtle and implicit measure than explicit measures but is less implicit than the SC-IAT-P, where participants appear to engage in a lower-level processing.

From a methodological perspective, LIB would be all the more relevant in measuring a bias if their assessment were computerized, under time constraints, and presented in the form of Likert-type scales focusing on positive items. Indeed, the results demonstrate that negative items regarding Maghrebians overtly express negative biases, leading to non-response reactions. When participants accept to respond to negative items, their scores are less strongly associated with the SC-IAT-P score than positive items. On the other hand, they are more strongly associated with the explicit measure of prejudice than with positive items.

In future research, it would be important to confirm that LIB can be utilized as an analytical tool for language productions as an observable of implicit prejudice regarding political orientation. This confirmation, in our view, should be based on positive items to be evaluated on a Likert-type scale, providing more variability in response possibilities. The study of abstraction levels in language as a means of assessing bias appears more suitable as a measure of explicit attitudes and can be directly applied to a corpus or spontaneous discourse. Thus, it would be possible to analyze free discourse, even if challenging from the perspective of LIBs (Dragogevic et al., 2017; Menegatti & Rubini 2017; see Rubin et al., 2014 for a review). This approach allows for significant freedom in studying phenomena related to racism and discrimination due to its straightforward implementation compared to other measures of implicit or indirect prejudices.

Thus, beyond the previously mentioned limitations, the major contribution of this research is to demonstrate that the level of language abstraction is indeed the result of implicit associations in

low-level cognitive memory, which are not observable and of which individuals may not necessarily be aware or have access to through consciousness.

References

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality*. Harpers.

Assilaméhou-Kunz, Y., Postmes, T., & Testé, B. (2020). A normative perspective on the linguistic intergroup bias: How intragroup approval of ingroup members who use the linguistic intergroup bias perpetuates explicit intergroup bias. *European Journal of Social Psychology, 50*(1), 81–96. <https://doi.org/10.1002/ejsp.2616>

Azar, O. H. (2008). Evolution of social norms with heterogeneous preferences: A general model and an application to the academic review process. *Journal of Economic Behavior & Organization, 65*(3-4), 420-435. <https://doi.org/10.1016/j.jebo.2006.03.006>

Bardin, B., Perrissol, S., Py, J., Launay, C., & Escoubès, F. (2014). Personalized SC-IAT: a possible way of reducing the influence of societal views on assessments of implicit attitude toward smoking. *Psychological reports, 115*(1), 13–25. <https://doi.org/10.2466/18.07.PR0.115c10z8>

Bardin, B., Perrissol, S., Py, J., Fos, Y., & Souchon, N. (2016). Testing of a paper-and-pencil Personalized Single Category Implicit Association Test (SC-IAT-P). *International Review of Social Psychology, 29*(1), 31-44. <https://doi.org/10.5334/irsp.35>

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.

<https://doi.org/10.1037/0022-3514.62.6.893>

Beauchemin, C., Hamel, C., Lesné, M., Simon, P., & l'équipe de l'enquête TeO. (2010). Les discriminations : une question de minorités visibles. *Population & Sociétés*, 466, 1-4.

<https://doi.org/10.3917/popsoc.466.0001>

Berezin, M. (2009). *Illiberal politics in neoliberal times: culture, security and populism in the new Europe*. Cambridge: Cambridge University Press.

Bierbrauer, G., & Klinger, E. W. (2002). Political ideology, perceived threat, and justice towards immigrants. *Social Justice Research*, 15(1), 41–52.

<https://doi.org/10.1023/A:1016045731732>

Bos, L., Sheets, P., & Boomgaarden, H. G. (2018). The Role of Implicit Attitudes in Populist Radical-Right Support. *Political Psychology*, 39(1), 69–87.

<http://www.jstor.org/stable/45094417>

Burguet, A. (2022). Mesure de l'attitude explicite et implicite à l'égard des migrants : le rôle de l'orientation politique. In Agnoletti M.F. & Salès-Wuillemin, E.(Eds), *Communications réelles et virtuelles : Nouvelles perspectives en psychologie sociale de la communication du langage*. Dijon: Editions universitaires de Dijon.

Campbell, D., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Daniels, J. P., & Von Der Ruhr, M. (2003). The determinants of immigration-policy preferences in advanced economies: A cross-country study. *Atlantic Economic Journal*, 31, 146-

158. <https://doi.org/10.1007/BF02319867>

Dasgupta, N., McGhee, D. E., Greenwald, A. G., & Banaji, M. R. (2000). Automatic preference for White Americans: Eliminating the familiarity explanation. *Journal of Experimental Social Psychology*, 36(3), 316-328. <https://doi.org/10.1006/jesp.1999.1418>

Douglas, K. M., & Sutton, R. M. (2006). When what you say about others says something about you: Language abstraction and inferences about describers' attitudes and goals. *Journal of Experimental Social Psychology*, 42(4), 500-508. <https://doi.org/10.1016/j.jesp.2005.06.001>

Douglas, K. M., Sutton, R. M., & Wilkin, K. (2008). Could you mind your language? An investigation of communicators' ability to inhibit linguistic bias. *Journal of Language and Social Psychology*, 27(2), 123-139. <https://doi.org/10.1177/0261927X07313655>

Douglas, K. M., & Sutton, R. M. (2010). By their words ye shall know them: Language abstraction and the likeability of describers. *European Journal of Social Psychology*, 40(2), 366-374. <https://doi.org/10.1002/ejsp.634>

Dragojevic, M., Giles, H., Beck, A.-C., & Tatum, N. T. (2017). The fluency principle: Why foreign accent strength negatively biases language attitudes. *Communication Monographs*, 84(3), 385-405. <https://doi.org/10.1080/03637751.2017.1322213>

Dragojevic, M., Sink, A., & Mastro, D. (2017). Evidence of Linguistic Intergroup Bias in U.S. Print News Coverage of Immigration. *Journal of Language and Social Psychology*, 36(4), 462-472. <https://doi.org/10.1177/0261927X16666884>

Esses, V. M., Dovidio, J. F., Jackson, L. M., & Armstrong, T. L. (2001). The immigration dilemma: The role of perceived group competition, ethnic prejudice, and national identity. *Journal of Social Issues*, 57(3), 389-412. <https://doi.org/10.1111/0022-4537.00220>

Esses, V. M., Jackson, L. M., Dovidio, J. F., & Hodson, G. (2005). Instrumental Relations Among Groups: Group Competition, Conflict, and Prejudice. In J. F. Dovidio, P. Glick, & L. A.

Rudman (Eds.), *On the nature of prejudice: Fifty years after Allport* (pp. 227–243). Blackwell Publishing. <https://doi.org/10.1002/9780470773963.ch14>

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. <https://doi.org/10.1037/0022-3514.50.2.229>

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

Feldman, S. (1988). Structure and consistency in public opinion: The role of core beliefs and values. *American Journal of Political Science*, *32*, 416–440. <http://dx.doi.org/10.2307/2111130>

Fitzpatrick, T. B. (1988). The validity and practicality of sun-reactive skin types I through VI. *Archives of dermatology*, *124*(6), 869–871. <http://dx.doi.org/10.1001/archderm.1988.01670060015008>

Franco, F. M., & Maass, A. (1996). Implicit versus explicit strategies of out-group discrimination: The role of intentional control in biased language use and reward allocation. *Journal of Language and Social Psychology*, *15*(3), 335–359. <https://doi.org/10.1177/0261927X960153007>

Gawronski, B. (2002). What does the Implicit Association Test measure? A test of the convergent and discriminant validity of prejudice-related IATs. *Experimental Psychology*, *49*(3), 171–180. <https://doi.org/10.1026/1618-3169.49.3.171>

Gawronski, B., Galdi, S., & Arcuri, L. (2015). What can political psychology learn from implicit measures? Empirical evidence and new directions. *Political Psychology, 36*(1), 1-17.
<https://doi.org/10.1111/pops.12094>

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*(1), 4–27. <https://doi.org/10.1037/0033-295X.102.1.4>

Greenwald, A. G., Draine, S. C., & Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science, 273*(5282), 1699-1702.
<http://dx.doi.org/10.1126/science.273.5282.1699>

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. <https://doi.org/10.1037/0022-3514.74.6.1464>

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

Hamberger, J., & Hewstone, M. (1997). Inter-ethnic contact as a predictor of blatant and subtle prejudice: Tests of a model in four West European nations. *British Journal of Social Psychology, 36*(2), 173–190. <https://doi.org/10.1111/j.2044-8309.1997.tb01126.x>

Hayes, A. F. (2018). Partial, conditional, and moderated moderated mediation: Quantification, inference, and interpretation. *Communication Monographs, 85*(1), 4–40.
<https://doi.org/10.1080/03637751.2017.1352100>

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.

<https://doi.org/10.1177/0146167205275613>

Huber, G. A., & Lapinski, J. S. (2006). The “race card” revisited: Assessing racial priming in policy contests. *American Journal of Political Science*, 50(2), 421-440.

<https://doi.org/10.1111/j.1540-5907.2006.00192.x>

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

<https://doi.org/10.1037/0022-3514.91.1.16>

Kaskeleviciute, R., Knupfer, H., & Matthes, J. (2023) Who Says “Muslims Are Not Terrorists”? News Differentiation, Muslim versus Non-Muslim Sources, and Attitudes toward Muslims, *Mass Communication and Society*. <https://doi.org/10.1080/15205436.2023.2268097>

[Hamberger.org/10.1080/15205436.2023.2268097](https://doi.org/10.1080/15205436.2023.2268097)

Kessler, A. E., & Freeman, G. P. (2005). Public Opinion in the EU on Immigration from Outside the Community. *JCMS: Journal of Common Market Studies*, 43(4), 825-850.

<https://doi.org/10.1111/j.1468-5965.2005.00598.x>

Lapeyronnie, D. (2008). *Ghetto urbain. Ségrégation, violence, pauvreté en France aujourd'hui*. Paris, Robert Laffont. <https://doi.org/10.4000/lectures.5542>

Li, Y., & Hills, T. T. (2021). Language patterns of outgroup prejudice. *Cognition*, 215, 104813. <https://doi.org/10.1016/j.cognition.2021.104813>

Maass, A. (1999). Linguistic intergroup bias: Stereotype perpetuation through language. In M. P. Zanna (Ed.), *Advances in experimental social psychology*, Vol. 31, pp. 79–121). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60272-5](https://doi.org/10.1016/S0065-2601(08)60272-5)

Maass, A., Salvi, D., Arcuri, L., & Semin, G. R. (1989). Language use in intergroup contexts: The linguistic intergroup bias. *Journal of personality and social psychology*, *57*(6), 981. <https://doi.org/10.1037//0022-3514.57.6.981>

Maass, A., Milesi, A., Zabbini, S., & Stahlberg, D. (1995). Linguistic intergroup bias: differential expectancies or in-group protection? *Journal of personality and social psychology*, *68*(1), 116–126. <https://doi.org/10.1037//0022-3514.68.1.116>

Maass, A., Ceccarelli, R., & Rudin, S. (1996). Linguistic intergroup bias: Evidence for in-group-protective motivation. *Journal of Personality and Social Psychology*, *71*(3), 512. <https://doi.org/10.1037/0022-3514.71.3.512>

Maier, M., Welzenbach-Vogel, I.C., Christner, C., Tillman, E.R., Zinkernagel, A. & Schmitt, M. (2023). Implicit and explicit populist and anti-immigrant attitudes and their explanatory power for populist radical-right party support. *Acta Politica*, *58*, 591–613. <https://doi.org/10.1057/s41269-022-00255-6>

Malchow-Møller, N., Munch, J. R., Schroll, S., & Skaksen, J. R. (2008). Attitudes towards immigration—Perceived consequences and economic self-interest. *Economics Letters*, *100*(2), 254–257. <https://doi.org/10.1016/j.econlet.2008.02.003>

Mendiburo-Seguel, A., & Ford, T. E. (2019). The effect of disparagement humor on the acceptability of prejudice. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues*. Advance online publication. <https://doi.org/10.1007/s12144-019-00354-2>

Mastro, D., Tukachinsky, R., Behm-Morawitz, E., & Blecha, E. (2014). News coverage of immigration: The influence of exposure to linguistic bias in the news on consumer's racial/ethnic cognitions. *Communication Quarterly*, *62*(2), 135-154. <https://doi.org/10.1080/01463373.2014.890115>

Matthes, J., & Marquart, F. (2015). A new look at campaign advertising and political engagement: Exploring the effects of opinion-congruent and -incongruent political advertisements. *Communication Research*, 42(1), 134–155. <https://doi.org/10.1177/0093650213514600>

Matthes, J., & Schmuck, D. (2015). The effects of anti-immigrant right-wing populist ads on implicit and explicit attitudes: A moderated mediation model. *Communication Research*, 44(4), 556-581. <https://doi.org/10.1177/0093650215577859>

Mayda, A. M. (2006). Who is against immigration? A cross-country investigation of individual attitudes toward immigrants. *The review of Economics and Statistics*, 88(3), 510-530. <https://doi.org/10.1162/rest.88.3.510>

McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology*, 37(5), 435–442. <https://doi.org/10.1006/jesp.2000.1470>

Menegatti, M., & Rubini, M. (2009). Le parole delle relazioni sociali: due decadi con il modello delle categorie linguistiche. *Psicologia sociale*, 4(3), 407-436. <https://doi.org/10.1177/0146167213479404>

Menegatti, M., & Rubini, M. (2013). Convincing similar and dissimilar others: The power of language abstraction in political communication. *Personality and Social Psychology Bulletin*, 39(5), 596-607. <https://doi.org/10.1177/0146167213479404>

Menegatti, M., & Rubini, M. (2017). Gender Bias and Sexism in Language. Oxford Research Encyclopedia of Communication. <https://doi.org/10.1093/acrefore/9780190228613.013.470>

Miguet, F. (2008). Voting about immigration policy: What does the Swiss experience tell us?. *European Journal of Political Economy*, 24(3), 628-641. <https://doi.org/10.1016/j.ejpoleco.2008.06.005>

Nisbett, R. E., & Bellows, N. (1977). Verbal reports about causal influences on social judgments: Private access versus public theories. *Journal of Personality and Social Psychology*, 35(9), 613–624. <https://doi.org/10.1037/0022-3514.35.9.613>

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

Nosek, B. A., Graham, J., & Hawkins, C. B. (2010). Implicit political cognition. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 548–564). The Guilford Press.

Olsson, F. (2023). The effect of implicit racial bias on right-wing populist support. *French Politics*, 21(1), 81-103. <https://doi.org/10.1057/s41253-022-00201-0>

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, 105(2), 171–192. <https://doi.org/10.1037/a0032734>

Paulhus, D. L. (1998). Interpersonal and intrapsychic adaptiveness of trait self-enhancement: A mixed blessing? *Journal of Personality and Social Psychology*, 74(5), 1197–1208. <https://doi.org/10.1037/0022-3514.74.5.1197>

Perugini, M., Richetin, J., & Zogmaister, C. (2010). Prediction of behavior. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 255–277). The Guilford Press.

Pettigrew, T. F., & Meertens, R. W. (1995). Subtle and blatant prejudice in western Europe. *European Journal of Social Psychology*, 25(1), 57–75.

<https://doi.org/10.1002/ejsp.2420250106>

Rubini, M., Menegatti, M., & Moscatelli, S. (2014). The strategic role of language abstraction in achieving symbolic and practical goals. *European Review of Social Psychology*, 25(1), 263–313. <https://doi.org/10.1080/10463283.2014.985501>

Rudman, L. A., Greenwald, A. G., Mellott, D. S., & Schwartz, J. L. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social cognition*, 17(4), 437-465. DOI: [10.1521/soco.1999.17.4.437](https://doi.org/10.1521/soco.1999.17.4.437)

Scaffidi Abbate, C., Giammusso, I., & Boca, S. (2020). The effect of perspective-taking on linguistic intergroup bias. *Journal of Language and Social Psychology*, 39(2), 183-199.

<https://doi.org/10.1177/0261927X19874383>

Schimmack (2019). The Implicit Association Test: A method in search of a construct. *Perspectives on Psychological Science*, <https://doi.org/10.1177/1745691619863798>

Schnake, S. B., & Ruscher, J. B. (1998). Modern racism as a predictor of the linguistic intergroup bias. *Journal of Language and Social Psychology*, 17(4), 484–491.

<https://doi.org/10.1177/0261927X980174004>

Schüller, S. (2016). The effects of 9/11 on attitudes toward immigration and the moderating role of education. *Kyklos*, 69(4), 604–632. <https://doi.org/10.1111/kykl.12122>

Semin, G. R. (2012). The linguistic category model. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 309–326). Sage Publications Ltd. <https://doi.org/10.4135/9781446249215.n16>

Semin, G. R., & Fiedler, K. (1988). The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of personality and Social Psychology*, 54(4), 558. <https://doi.org/10.1037/0022-3514.54.4.558><https://doi.org/10.1037/0022-3514.54.4.558>

Semin, G.R., & Fiedler, K. (1989). Relocating attributional phenomena within a language-cognition interface: The case of actors' and observers' perspectives. *Eur. J. Soc. Psychol.*, 19: 491-508. <https://doi.org/10.1002/ejsp.2420190602>[s://doi.org/10.1002/ejsp.2420190602](https://doi.org/10.1002/ejsp.2420190602)

Semin, G. R., & Fiedler, K. (1991). The Linguistic Category Model, its Bases, Applications and Range. *European Review of Social Psychology*, 2(1), 1–30. <https://doi.org/10.1080/14792779143000006>

Smith, C. T., De Houwer, J., & Nosek, B. A. (2013). Consider the source: Persuasion of implicit evaluations is moderated by source credibility. *Personality and Social Psychology Bulletin*, 39(2), 193-205. <https://doi.org/10.1177/0146167212472374>

Snefjella, B., & Kuperman, V. (2015). Concreteness and psychological distance in natural language use. *Psychological Science*, 26(9), 1449–1460. <https://doi.org/10.1177/0956797615591771>

Stephan, W. G., Stephan, C. W., & Gudykunst, W. B. (1999). Anxiety in intergroup relations: A comparison of anxiety/uncertainty management theory and integrated threat theory. *International Journal of Intercultural Relations*, 23(4), 613–628. [https://doi.org/10.1016/S0147-1767\(99\)00012-7](https://doi.org/10.1016/S0147-1767(99)00012-7)

Tournois, J., Mesnil, F., & Kop, J. L. (2000). Self-Deception and Other-Deception: A Social Desirability Questionnaire. *European Review of Applied Psychology*, 50(1), 219-233.

Voci, A., & Hewstone, M. (2003). Intergroup contact and prejudice toward immigrants in Italy: The mediational role of anxiety and the moderational role of group salience. *Group Processes & Intergroup Relations*, 6(1), 37-54. <https://doi.org/10.1177/1368430203006001011>

Von Hippel, W., Sakaquaptewa, D., & Vargas, P. (1997). The linguistic intergroup bias as an implicit indicator of prejudice. *Journal of Experimental Social Psychology*, 33(5), 490-509. <https://doi.org/10.1006/jesp.1997.1332>

Wakslak, C., Smith, P., & Han, A. (2014). Using abstract language signals power. *Journal of personality and social psychology*, 107(1), 41-55. <http://dx.doi.org/10.1037/a0036626>

Wigboldus, D. H. J., Semin, G. R., & Spears, R. (2000). How do we communicate stereotypes? Linguistic bases and inferential consequences. *Journal of Personality and Social Psychology*, 78(1), 5–18. <https://doi.org/10.1037/0022-3514.78.1.5>

Wigboldus, D., & Douglas, K. (2007). Language, Stereotypes, and Intergroup Relations. In K. Fiedler (Ed.), *Social communication* (pp. 79–106). Psychology Press.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, 72(2), 262–274. <https://doi.org/10.1037/0022-3514.72.2.262>