Received: 30 March 2018 | Accepted: 2 December 2019

DOI: 10.1002/ejsp.2653

RESEARCH ARTICLE

EASP WILEY

Toward the use of approach/avoidance tendencies as attitude measures: Individual- and group-level variability of the ingroup bias

Marine Rougier[1,2] | Dominique Muller[1,3] | Rémi Courset[1] | Annique Smeding[4] | Thierry Devos[5] | Cédric Batailler[1]

[1]LIP/PC2S, Univ. Grenoble Alpes, Grenoble, France

[2]IPSY, Université Catholique de Louvain, Louvain-la-Neuve, Belgique

[3]Institut Universitaire de France, Paris, France

[4]LIP/PC2S, Univ. Savoie Mont Blanc, Chambery, France

[5]San Diego State University, San Diego, CA, USA

Correspondence
Dominique Muller, LIP/PC2S, Univ. Grenoble Alpes, CS 40700, 38 058 Grenoble Cedex 9, France.
Email: dominique.muller@univ-grenoble-alpes.fr

Funding information
Agence Nationale de la Recherche, Grant/Award Number: ANR-15-IDEX-02 and ANR-2010-BLAN-1905-01; CNRS; Pôle Grenoble Cognition; SFR Santé & Société

Abstract

Approach/avoidance paradigms could constitute an interesting alternative in measuring intergroup attitudes, notably if they overcome one criticism often addressed toward classic indirect tasks: Measuring attitudes beyond the influence of cultural knowledge. Using intergroup stimuli and a population likely to be exposed to a similar cultural knowledge, we observed two informative results regarding this issue: Approach/avoidance effects measured by the Visual Approach/Avoidance by the Self Task (VAAST) varied across participants (i.e., consistent with the variability of intergroup attitudes; Experiment 1) and both participants of dominant and non-dominant groups produced an ingroup bias (Experiment 2). A last experiment (Experiment 3) showed that compatibility scores in the VAAST predict trustworthiness ratings of the ingroup/outgroup. This experiment also investigated potential differences between the VAAST and the IAT. These results suggest that approach/avoidance tasks (notably the VAAST) could be relevant to assess personal attitudes when it comes to normatively sensitive topics.

KEYWORDS

approach/avoidance measure, indirect measure, intergroup attitudes, VAAST

1 | INTRODUCTION

General positive versus negative inclinations toward stimuli (as social groups), or what social psychologists often call "attitudes" (Eagly & Chaiken, 1993), are often assessed with indirect measures. The advantage of these latter measures lies in the fact that they are often considered as less influenced by self-presentation concerns compared to direct measures (e.g., self-report, Fazio, Jackson, Dunton, & Williams, 1995). Beyond classic indirect tasks (e.g., Implicit Association Test [IAT], affective priming), research on indirect measures of attitudes is constantly expanding through the development of new measurement paradigms. Among them, measures focusing on different aspects of attitudes, such as approach/avoidance tendencies, constitute an interesting alternative to measure intergroup attitudes and represent an ever-expending research area. However, if approach/avoidance tasks are to win acclaim as relevant measures of intergroup attitudes, they have to successfully address a central criticism often expressed toward classic indirect measures, namely, whether approach/avoidance tasks could discriminate individuals on their (personal) intergroup attitudes beyond the influence of cultural knowledge. First, we reasoned that if approach/avoidance measures can operate such discrimination, effects should vary across individuals and between social groups, even if these individuals and groups are exposed to a similar cultural knowledge. Second, the approach/avoidance effects should be linked to a criterion variable correlating with indirect measures beyond direct measures of prejudice. We tested these questions in three experiments and in the last experiment we also investigated differences/similarities with a classic indirect measure (i.e., IAT).

1.1 | Indirect measures of attitudes and the cultural knowledge issue

Indirect measures of attitudes toward social groups often rely on individuals' performances, typically through a limited response time setting (Gawronski & De Houwer, 2014). In these tasks, designed to capture intergroup attitudes, as in the Race IAT (Greenwald, McGhee, & Schwartz, 1998) and the affective priming paradigm (Fazio et al., 1995), European Americans display on average a response pattern indicative of intergroup bias. Precisely, bias is inferred from the finding that, on average, European Americans associate more easily positive versus negative stimuli with typical European American versus African American stimuli, respectively. In the last decades, in addition to the development of many additional indirect measures of attitudes (e.g., First-Person Shooter Task; Correll, Park, Judd, & Wittenbrink, 2002; Extrinsic Affective Simon Task; De Houwer, 2003; Go/No-Go Association Task; Nosek & Banaji, 2001; Affect Misattribution Procedure [AMP]; Payne, Cheng, Govorun, & Stewart, 2005), several approach/avoidance paradigms started to be developed (e.g., Lever task; Chen & Bargh, 1999; Manikin task; De Houwer, Crombez, Baeyens, & Hermans, 2001; Joystick task; Rinck & Becker, 2007; Modified keyboard; Vaes, Paladino, Castelli, Leyens, & Giovanazzi, 2003). The rationale behind these latter studies is that faster response time to approach (vs. avoid) a given stimulus (e.g., an ingroup first name) would represent an approach (vs. avoidance) tendency toward this stimulus. Ultimately, assessing approach/avoidance tendencies toward a stimulus would be informative of the general attitude toward it.

By using a modified keyboard, Paladino and Castelli (2008) showed several approach/avoidance intergroup effects, with participants—only members of the dominant group (i.e., groups associated with greater social value, such as white people in most Western societies)—being faster in the compatible setting (i.e., approach ingroup stimuli and avoid outgroup stimuli by moving one's hand toward vs. away from the screen to push keyboard buttons) as compared to the incompatible setting where the approach/avoidance instructions were reversed (see also Bianchi, Carnaghi, & Shamloo, 2018; Castelli, Zogmaister, Smith, & Arcuri, 2004; Clow & Olson, 2010; Degner, Essien, & Reichardt, 2016; Neumann, Hülsenbeck, & Seibt, 2004; Vaes et al., 2003). So far, however, these approach/avoidance measures have not often been expressly used as attitude measures.

At the conceptual level, however, approach/avoidance tendencies seem to meet the defining criteria of an attitude. According to some authors, the definition of attitudes has indeed evolved from a "neural state of readiness" (Allport, 1935, p. 810) to a definition focusing more on the approach/avoidance consequences associated with an "evaluative predisposition" (Krosnick, Judd, & Wittenbrink, 2005). At the empirical level, several experiments were able to show a link between evaluative stimuli and behavioral tendencies: Positive stimuli would automatically trigger approach tendencies and negative stimuli avoidance tendencies (e.g., Chen & Bargh, 1999; De Houwer et al., 2001; Markman & Brendl, 2005; Rinck & Becker, 2007; Rougier et al., 2018; Seibt, Neumann, Nussinson, & Strack, 2008; Wentura, Rothermund, & Bak, 2000). As for other kinds of indirect measures, response times are here paramount, with the response time difference between compatible and incompatible blocks/trials supposedly indicative of an individual's personal attitudes (Cacioppo, Priester, & Berntson, 1993; Chen & Bargh, 1999; Neumann et al., 2004). A major challenge, however, if one is to investigate individuals' personal and (often private) attitudes with indirect measures, is determining to what extent these measures are subject to what could be called the "cultural knowledge issue".

In the present contribution, and in line notably with Payne, Vuletich, and Lundberg (2017), we define cultural knowledge as the objective exposure to a given environment, spreading certain ideas and values about social groups (e.g., "North African people are bad people"). This definition can be also called the "concept accessibility" shared by individuals belonging to the same culture (e.g., Payne et al., 2017; see also Lynott, Kansal, Connell, & O'Brien, 2012), or "extra-personal associations" (Han, Olson, & Fazio, 2006; Karpinski & Hilton, 2001).[1] Following this idea, because cultural exposure can come from multiple sources (e.g., traditional or social media, sports club, etc.), cultural knowledge could theoretically be defined across as many levels as there are sources of cultural influence. Yet in this work, and as developed thereafter, we chose to rely on the geographical area because individuals coming from the same geographical area would be chronically exposed to a similar cultural knowledge (Kurdi & Banaji, 2017; Shepherd, 2011) and because this level has been shown to be informative regarding the cultural knowledge issue. On the contrary, and in line with classic definitions, we define the personal attitudes as "neural states of readiness" (Allport, 1935) coming from "traces of our past experiences" (Greenwald & Banaji, 1995) with stimuli. In other words, if two persons differ in states of readiness to behave in a prejudiced fashion, they have different personal attitudes.

According to some authors, because classic indirect measures are sensitive to the environmental context in which they are completed (Han et al., 2006; Karpinski & Hilton, 2001; Lowery, Hardin, & Sinclair, 2001; Lynott et al., 2012), they would merely capture the dominant cultural message that is prevalent in a given society, but not the inter-individual variability that may exist among individuals regarding a given attitude object (Arkes & Tetlock, 2004; Olson & Fazio, 2004; Payne et al., 2017). Arkes and Tetlock (2004) even argued that "If I am aware of the cultural stereotype, I have all the cognitive software that I need to manifest prejudice on the IAT" (p. 262). Following this extreme position, it may be argued that classic indirect measures struggle to discriminate individuals who are truly prejudiced (i.e., whose personal attitudes are in line with the prevalent cultural knowledge) from individuals who are simply aware of this knowledge (i.e., individuals exposed to the same cultural environment, but having different personal

[1] It should be noted that, here, we do not conceive cultural knowledge in the same way as other authors do (e.g., Nosek & Hansen, 2008), that is, as the subjective perception of prejudice toward social groups, but rather as the objective exposure to a biased information (e.g., extra-personal associations) about these groups.

attitudes; Arkes & Tetlock, 2004; see also Gawronski & Bodenhausen, 2006, 2007, 2014; Strack & Deutsch, 2004). Relatedly, in a recent debate, Payne et al. (2017) proposed that classic indirect measures (i.e., the IAT and the AMP) would be more effective in measuring situational biases (e.g., cultural bias) than individual biases (Payne et al., 2017; see the Section 5 for more details).

Even if some authors clearly endorse these extreme views, we want to insist that we do not imply that classic indirect measures of attitudes either tap only into cultural knowledge or into personal prejudice. In fact, empirical results are not conclusive regarding these possibilities. On the one hand, meta-analyses show that these measures are able to capture a personal content. For instance, IAT scores are significantly linked with direct measures (Greenwald, Poehlman, Uhlmann, & Banaji, 2009; Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005), behaviors (Kurdi et al., 2019; but see Oswald, Mitchell, Blanton, Jaccard, & Tetlock, 2013), and brain activity (Oswald et al., 2013). Moreover, the relationship between IAT scores and direct measures of prejudice seems to go beyond the awareness of a cultural bias (Nosek & Hansen, 2008; but see Footnote 1). On the other hand, two kinds of results are often considered as a limitation of classic indirect tasks in measuring personal prejudice.

First, among members of the dominant group (e.g., white people), effects produced by classic indirect measures generally show a surprisingly high (Correll et al., 2002; Fazio et al., 1995; Greenwald et al., 1998; Nosek, Banaji, & Greenwald, 2002) and homogeneous (Judd, Westfall, & Kenny, 2012) bias in favor of the dominant group. Interestingly for our purpose, Judd et al. (2012) reanalyzed the data of Correll et al. (2002) with a mixed model technique testing whether there was significant variability of the shooter bias across participants (i.e., what we called here "individual-level variability"). Specifically, this kind of analysis tests whether participants display either variability in the amount of bias, with stronger/weaker bias for some individuals than for others (in which case individual-level variability would be significant), or a similar level of bias (in which case this variability would not be significant). Perhaps surprisingly, they found no such significant variability, meaning that individuals had a similar shooter bias. Importantly, this effect was found in a population likely to be exposed to a similar cultural knowledge (i.e., students from the same campus). Second, previous work has often found that such pro-dominant group bias is also present among individuals belonging to non-dominant social groups (e.g., non-white people; Ashburn-Nardo, Knowles, & Monteith, 2003; Nosek et al., 2002; Richeson, Trawalter, & Shelton, 2005). Using the Race IAT, Nosek et al. (2002) notably showed, through a very large sample, that African Americans present on average a pro-European American bias. These results are consistent with the idea that the social value granted to social groups in a given social context has a large influence on the effects produced by indirect measures of attitudes (Jost, Pelham, & Carvallo, 2002; Rudman, Feinberg, & Fairchild, 2002).

Therefore, based on the literature, it seems that classic indirect attitude measures are often not well suited to capture variability in intergroup attitude, be it (a) across individuals—the "individual-level variability"—or (b) across social groups—what we call "group-level variability" (e.g., reversed bias for some social groups). In this work, we leave aside the debate on classic indirect measures' ability to capture or not personal prejudice. Indeed, our focus here is on whether the less classic indirect measures of approach/avoidance tendencies, in particular the Visual Approach/Avoidance by the Self Task (VAAST; Rougier et al., 2018), are suitable for capturing individual-level and group-level variability. Accordingly, we tested whether these last two limitations—in discriminating between individuals and groups beyond cultural knowledge—are likely to apply to or, instead, may be overcome by approach/avoidance measures.

1.2 | Are approach/avoidance effects likely to be affected by the cultural knowledge issue?

What differentiates approach/avoidance measures from classic indirect measures is the dimension of the attitude they assess (i.e., approach/avoidance tendencies toward social groups vs. positive/negative associations with social groups) and its conceptual relationship with cultural knowledge (see also Lynott et al., 2012 for a similar reasoning). Although attitudes toward social groups can be both defined through positive/negative and approach/avoidance aspects, as we argued previously, the latter aspect would possess a special connection with the self—and therefore could be less influenced by cultural knowledge. Crucially, because approach/avoidance actions are self-generated bodily reactions implemented in everyday life, the approach/avoidance dimension is more likely to pertain to the individual rather than to the culture. Of course, there is cultural knowledge about which group should be approached or avoided (e.g., "North Africans should be avoided"); however, the everyday life implementation of approach/avoidance behaviors passes through the body. Accordingly, a task reproducing these real-life bodily activations (as the VAAST does) should be more likely to capture individual reactions rather than cultural knowledge. The positive/negative dimension, conversely, is common to both attitudes (i.e., positive/negative evaluation of social groups made by the self) and cultural knowledge (e.g., positive/negative value granted to social groups). This last point is crucial: While the positive/negative aspects associated with social groups could come from both attitudes and environment, approach/avoidance aspects should be more self-referenced. Accordingly, approach/avoidance tendencies could potentially capture individual-level and group-level variability beyond the influence of cultural knowledge about social groups.

Although limited, recent work on approach/avoidance is consistent with the idea that these effects could be effective in capturing an individual's personal attitude. First, several studies showed that approach/avoidance effects depend on characteristics related to personal experiences (e.g., Clow & Olson, 2010; Cousijn, Goudriaan, & Wiers, 2011; Rinck & Becker, 2007; Wiers et al., 2013; Zhou et al., 2012). For instance, faster response time to avoid spider pictures and to approach non-spider pictures rather than the reverse depends on the individual's phobia scores toward spiders (Rinck & Becker, 2007). Second, at the group-level, Paladino and Castelli (2008) were able

to show approach/avoidance effects using the minimal group paradigm (i.e., "yellow" group vs. "red" group): Participants were faster to approach ingroup members and to avoid outgroup members rather than the reverse. In this last experiment, however, groups were fictitious and therefore not subject to a potential influence of cultural knowledge. Overall, at the time we conducted this research, no previous experiments using real social groups had assessed, using approach/avoidance measures, what we defined as individual- and group-level variability (but see Bianchi et al., 2018; Degner et al., 2016). Specifically, no experiment tested (a) whether the pro-ingroup approach/avoidance compatibility effect is variable across individuals and (b) whether a pro-ingroup effect occurs in a full ingroup–outgroup design (i.e., with both dominant and non-dominant social groups producing evaluations of the two groups; Judd & Park, 1993). These two points leave open the question of whether these tendencies are robust to the cultural knowledge issue.

Accordingly, one needs to address whether approach/avoidance effects toward groups associated with a dominant and widespread cultural knowledge (such as ethnic groups) vary among individuals and depend on real group membership. In other words, can approach/avoidance effects capture individual variations for a given attitude and go beyond cultural knowledge? If this is the case, approach/avoidance measures could constitute an alternative to the widely used classic indirect measures of attitudes, often criticized on this point. The aim of our experiments was precisely to test whether approach/avoidance tasks were able to discriminate between individuals and between group memberships on racial intergroup attitudes.

1.3 | The present research

To address these two points, we ran three experiments. Experiment 1 tested the approach/avoidance variability at the individual level and Experiment 2 at both the individual- and group-levels. In a final experiment we investigated whether approach/avoidance tendencies could predict a measure known to correlate with an indirect measure of attitudes.

More specifically, Experiment 1 addressed the question of whether approach/avoidance effects toward racial social groups are variable across individuals by using two approach/avoidance tasks (i.e., the VAAST and Manikin task) where participants had to categorize first names according to their origin. Observing inter-individual variability is necessary but not sufficient to reach a firm conclusion on whether approach/avoidance tasks measure personal attitude. Indeed, not finding any variability can be indicative of a bias reflecting a shared (i.e., not variable) knowledge, but a significant variability of the bias could be due to other factors than variability in personal attitude (e.g., minimal differences in cultural exposure, executive functions, etc.). Accordingly, we also started to investigate whether the inter-individual variability was meaningfully explained by an individual-level variable (for a similar reasoning, see Fazio et al., 1995). To do so, we tested whether the inter-individual variability could be (at least partly) explained by a self-report measure of prejudice. Experiment 2 replicated the compatibility effect observed in Experiment 1 and tested the group-level variability of this effect, that is, the moderation of approach/avoidance effects toward social groups by group membership. Specifically, we tested whether the compatibility effect was pro-ingroup in both dominant and non-dominant groups (i.e., French origin and North African origin participants). Such findings would stand in contrast with classic indirect tasks, often showing a bias in favor of the dominant group, independently of group membership (Ashburn-Nardo et al., 2003; Nosek et al., 2002; Richeson et al., 2005). Finally, in Experiment 3 we tested whether the compatibility effect in the VAAST could predict scores obtained from ingroup/outgroup trustworthiness ratings. Indeed, showing in Experiment 1 that approach/avoidance tendencies are linked with a direct measure of prejudice can be somewhat problematic for theoretical approaches predicting a dissociation between explicit and implicit processes (e.g., Strack & Deutsch, 2004). Therefore, Experiment 3 investigated the relationship between approach/avoidance tendencies and a criterion variable known to correlate with an indirect measure (i.e., the IAT), above and beyond direct measures of attitudes: Trustworthiness ratings of ingroup/outgroup faces (Stanley, Sokol-Hessner, Banaji, & Phelps, 2011). In this experiment, we also investigated potential differences between approach/avoidance tendencies and IAT *D* score in predicting trustworthiness ratings.

As depicted beforehand, we define cultural knowledge relative to the geographical area: People coming from the same geographical area should be exposed to a similar cultural knowledge. Homogenizing a sample on the geographical area would thus favor the emergence of cultural knowledge issues, if any (i.e., a lack of inter-individual variability, Judd et al., 2012; or inter-group variability, Nosek et al., 2002). Accordingly, predictions in Experiments 1 and 2 were tested in a relatively homogeneous population (i.e., students of the same university), that is, in a population likely to be exposed to similar cultural knowledge. By doing so, the inter-individual and inter-group differences in approach/avoidance effects would be less likely to be due to differences in cultural exposure. Importantly, a classic indirect measure of attitudes did show a non-significant inter-individual variability (e.g., shooter bias; Judd et al., 2012) with the same kind of population (university students) and the same analytical method (mixed models) as ours. Conversely, Experiment 3 targeted a broader population in an online setting and among different social groups (i.e., African Americans/European Americans instead of French/North Africans as in Experiments 1 and 2). Because these participants come from all over the US, they also come with various cultural backgrounds. Indeed, even if we only recruited European Americans living in the US, participants were likely to come from states/cities having a non-homogeneous cultural knowledge about African versus European Americans. In line with this idea, we know from previous work that there is meaningful variability in the racial bias due to cities and counties (Hehman, Flake, & Calanchini, 2018; Orchard & Price, 2017; Payne, Vuletich, & Brown-Iannuzzi, 2019; Zerhouni,

Rougier, & Muller, 2016). Any inter-individual variability (or differences in inter-individual variability between the VAAST and the IAT) could thus be explained by a variability in cultural exposure. Accordingly, this last experiment was not suited to address the variability hypothesis. Nevertheless, because variability could still be of interest to some researchers, we added this information.

Regarding open practices, we provide the material, data, and data analysis (R scripts) from all experiments (link available at the end of the manuscript). Additionally, Experiment 3 was pre-registered on Open Science Framework (OSF). The pre-registration includes the a priori theoretical reasoning, hypotheses, power estimations, procedure, and analytical strategies.

2 | EXPERIMENT 1

The main goal of this experiment was to test whether approach/avoidance tasks could produce intergroup effects varying across participants. We chose the VAAST because Rougier et al. (2018) showed that this task produced strong and reliable compatibility effects and we used the Manikin task because Krieglmeyer and Deutsch (2010) found that this task produced larger compatibility effects than the famous Joystick task. Beyond producing large compatibility effects and contrary to other approach/avoidance tasks (e.g., Joystick task), these tasks are also unambiguous regarding the interpretation of the approach/avoidance actions they implement (Rougier et al., 2018; Seibt et al., 2008; and see the Section 5). Overall, we reasoned that the VAAST and the Manikin task's sensitivity should enable us to capture the variability of the compatibility effect among individuals, even if these individuals are homogeneously exposed to a given cultural knowledge about the ethnic groups of interest. We were also interested in whether this variability was linked with self-reported prejudice toward the outgroup (here, North African individuals). To our knowledge, two studies were able to show a link between intergroup approach/avoidance effects and self-reported prejudice (Clow & Olson, 2010; Degner et al., 2016), but such a link is not always observed (Neumann et al., 2004).

2.1 | Method

2.1.1 | Participants and design

To estimate our sample size for sufficient power (80%), we relied on Paladino and Castelli's compatibility effect sizes (2008; Studies 1a–1c, $d = 0.96$[2] on average), but also on our own past experiments with valenced stimuli using the VAAST (Rougier et al., 2018, $dz = 0.86$). We estimated that to reach sufficient power we needed at least 45 participants and we ended-up with 49 participants ($M_{age} = 20.64$, $SD_{age} = 2.29$, 36 female participants) who took part in the experiment in exchange for course credits—we removed 2 participants erroneously coded with the same number from our initial pool of 51 participants. Participants came from the same university. In this experiment, to analyze results for both VAAST and Manikin tasks, we used a 2 (compatibility: compatible vs. incompatible) × continuous (self-reported prejudice score: from 1 to 7) × 2 (block order: compatible first vs. incompatible first) × 2 (task order: VAAST first vs. Manikin task first) design with the last three variables being between participants.

For both tasks, participants went through a compatible block (i.e., approaching French first names and avoiding North African first names) and an incompatible block (i.e., approaching North African first names and avoiding French first names). We counterbalanced task and block orders between participants, block order being the same in each task for a specific participant. Each of the 40 first names (20 French and 20 North African) was randomly presented twice within each block of the two tasks, so that each block comprises 80 trials. Before each block, participants performed a training phase consisting in 8 trials over 4 first names that were not presented in the main experiment.

French and North African first names used for the approach/avoidance tasks came from the "Lexique" database (New, Pallier, Brysbaert, & Ferrand, 2004). We controlled the first names' frequency in order to create two groups of French first names: One group of 10 frequent first names ($M = 51.59$, $SD = 26.04$) and one group of 10 non-frequent first names ($M = 2.34$, $SD = 1.25$), this last one being equivalent to the group of 20 North African first names ($M = 2.21$, $SD = 2.07$). With these groups, we tested whether the compatibility effect depended on the ethnic origin of first names and/or their familiarity (e.g., faster response times to approach familiar first names and to avoid unfamiliar first names rather than the reverse; Jones, Young, & Claypool, 2011).[3]

2.1.2 | Procedure

The procedure was the same as Rougier et al. (2018, Experiment 1). Participants had to perform two approach/avoidance tasks. They performed the VAAST (Rougier et al., 2018) and the Manikin task (De Houwer et al., 2001) on a 23-inch computer screen (60 Hz). We used a chin rest to set the distance to the screen at 95 cm. For both tasks, participants categorized stimuli with a button box by using the index of their dominant hand. Three adjacent buttons were used: The middle button to start each trial and the other two to perform the categorization task. Participants had to keep their finger pressed on the start button until the word appeared on the screen and, when it appeared, to push one of the two end buttons on the button box depending on

[2]We estimated our sample size on the basis of a Cohen's *d* instead of a *dz* given that this one was not computable based on the available information from Paladino and Castelli (2008). Accordingly, our sample size estimation partly resulted from an approximate of *dz*, on the basis of Paladino and Castelli's average *d*.

[3]First names frequency did not significantly moderate approach/avoidance tendencies, either in the VAAST, $F(1, 48.95) = 0.04$, $p = .84$, or in the Manikin task, $F(1, 37.89) = 0.45$, $p = .50$; therefore, we removed this factor from the analysis.

the stimulus. Participants were encouraged to respond as quickly and as accurately as possible. After four key presses in the same direction (i.e., for a complete approach or avoidance movement), the trial terminated. For each trial, we recorded response times from the appearance of the word to the first push on one of the two categorization buttons.

2.1.3 | The VAAST

This task simulates approach/avoidance using visual information. Stimuli were displayed in a background giving an impression of depth and we displayed the back of the head of a person (see Figure 1). When participants pressed the start button, the white circle displayed in the center of the screen was replaced by a fixation cross (for a random duration of 800–2,000 ms), which was followed by a target first name. According to the participants' approach/avoidance action, the whole visual environment (i.e., the background image and the target word) was zoomed in (i.e., approach, "move forward" button) or zoomed out (i.e., avoidance, "move backward" button) by 10% for each button press (i.e., 0.13 angular degrees), giving the visual impression to walk forward or backward as a consequence of these actions. The stimuli, presented initially in font size 18 (Courier New typeface and white color given the dark background), could therefore vary from 30% larger (approach) to 30% smaller (avoidance).

2.1.4 | The Manikin task

This task had a white background and the manikin (a little schematic figure) as well as the font color were black. When participants pressed the start button, a fixation cross was displayed (for a random duration of 550–950 ms, following Krieglmeyer & Deutsch, 2010), followed by the manikin (displayed on the left or right side of the word) appearing 750 ms before the target name (displayed in the center of the screen). When the first name appeared on the screen, participants had to categorize it as being French or North African by moving the manikin (by 1.20 angular degrees) toward or away from the target word. The first names' size was always 32 (Arial typeface).

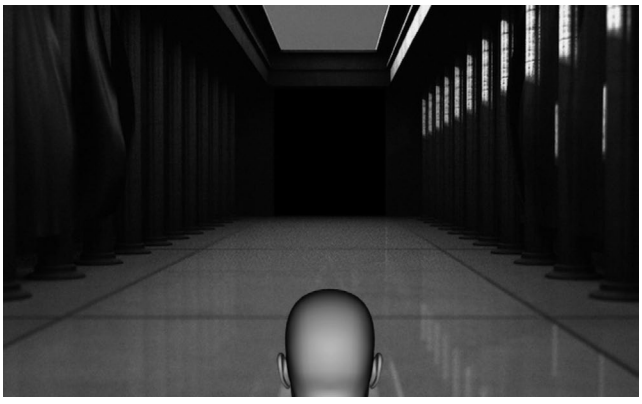

FIGURE 1 Background of the VAAST used in Experiment 1

After the approach/avoidance tasks, participants had to indicate whether they considered themselves to be of French, North African, or another origin. We chose, a priori, to keep only participants responding "French origin" (and we removed data from three participants who responded "North African origin" or "other origin"; see Experiment 2 for a test of the compatibility effect among participants categorizing themselves as being of North African origin). We also used two questions asking participants about their proficiency in the French language (one asking if French was their native language and, if it was not, a second one asking about their proficiency level). All participants reported having a high proficiency level.

2.1.5 | Self-report prejudice scale

At the end of the experiment, participants had to respond to a generalized prejudice scale toward North African individuals (Dambrun & Guimond, 2001). This scale was presented as an "Opinion scale about diversity and immigration in France". The experimenter insisted on the anonymity of the questionnaire and encouraged participants to answer spontaneously and honestly. Participants had to indicate their agreement from 1 (*not agree at all*) to 7 (*totally agree*) to 15 statements (e.g., "It is easy to understand the anger of North Africans in France", "French should come first when it comes to increase social benefits").

2.2 | Results

As main hypotheses, we expected a compatibility effect in both the VAAST and the Manikin task, but also that this compatibility effect would significantly vary across participants. To analyze response times (RTs) we removed incorrect trials (i.e., 3.33% of the trials for the VAAST and 4.77% of the trials for the Manikin task), as well as RTs faster than 400 ms and exceeding 1,700 ms (i.e., 2.61% of the trials for the VAAST and 4.54% of the trials for the Manikin task), and to normalize their distribution, we transformed RTs using an inverse function (Ratcliff, 1993). We selected these filters and transformations (out of several options) because they resulted in the most normal RT distribution—RT distributions and main results using other filters and transformations are presented as Tables S1–S5.

2.2.1 | Compatibility effect

To analyze our data, we used a mixed model analysis (Judd et al., 2012; Westfall, Kenny, & Judd, 2014). In this kind of analysis, and in contrast with traditional analyses of variance (e.g., ANOVA), multiple random factors are used (e.g., participants, stimuli) instead of one. Accordingly, mixed models allow us to generalize the results not only

over one but over every random term used in the model (i.e., over all participants and all stimuli at the same time) and thus maximize the generalizability of the findings compared to traditional analyses. Importantly for our purpose, this analysis also enables us to test the variability of an effect (e.g., compatibility effect) for a given random factor (e.g., participants). To the best of our knowledge, this important feature of mixed models has not yet been really applied in the context of indirect measures. Two exceptions being the shooter bias reanalysis mentioned in the introduction section (Judd et al., 2012) and a more recent proposition for adapting mixed models to indirect measures (Wolsiefer, Westfall, & Judd, 2016).

Because assessing differences in the VAAST and Manikin task was not the focus of this research, we analyzed the data for each of them separately. Accordingly, for each task, we estimated a model having the compatibility as fixed effect and we estimated random intercepts and the related slopes for participants, stimuli, and their interaction. The block and task orders did not moderate the compatibility effect, either in the VAAST, $F(1, 43.89) = 0.01$, $p = .93$, and, $F(1, 43.91) = 0.18$, $p = .67$, respectively, or in the Manikin task, $F(1, 44.25) = 0.02$, $p = .89$, and, $F(1, 44.27) = 0.16$, $p = .69$, respectively. Accordingly, we removed these control factors from the analysis.

In the VAAST, the compatibility effect was significant, $F(1, 51.99) = 5.60$, $p = .02$, $dz = 0.37$, indicating that participants were faster in the compatible block ($M = 697$ ms, $SE = 12$ ms) compared to the incompatible block ($M = 725$ ms, $SE = 14$ ms). Importantly, this compatibility effect varied significantly across participants, $\chi^2 = 135.33$, $p < .001$. In the Manikin task, participants were descriptively faster in the compatible block ($M = 728$ ms, $SE = 16$ ms) compared to the incompatible block ($M = 747$ ms, $SE = 18$ ms), but this compatibility effect was not significant, $F(1, 69.40) = 1.40$, $p = .24$, $dz = 0.23$. For the Manikin task, the compatibility effect also varied significantly across participants, $\chi^2 = 116.01$, $p < .001$. For both the VAAST and Manikin task, variability of the compatibility effect across stimuli and across the participants by stimuli interaction is presented in the Tables S1–S5.

Although the VAAST versus Manikin task comparison was not the focus of the present contribution, this question could still be of interest to other researchers. Accordingly, we tested whether the two tasks differed in producing a compatibility effect. To this aim, we estimated a model having the compatibility and the type of task (VAAST vs. Manikin task) as fixed effects and we estimated random intercepts and the related slopes for participants, stimuli, and their interaction. The overall compatibility effect was marginal, $F(1, 71.23) = 3.85$, $p = .054$, $dz = 0.37$, indicating that participants were marginally faster in the compatible block ($M = 712$ ms, $SE = 13$ ms) compared to the incompatible block ($M = 736$ ms, $SE = 15$ ms). Moreover, this effect was not significantly moderated by the type of task, $F(1, 58.00) = 0.38$, $p = .56$, $dz = 0.10$.

2.2.2 | Self-reported prejudice

We computed a prejudice score for each participant ($\alpha = .86$). To test whether variability of approach/avoidance tendencies was linked to self-reported prejudice, we conducted two analyses for both VAAST and Manikin task. One model tested whether variability of approach/avoidance tendencies toward North African first names was linked to self-reported prejudice as we though it should be. To do so, we estimated a mixed model while keeping only the response times related to North African first names and with the movement (i.e., approach vs. avoidance) as a fixed effect—that is, we decomposed the compatibility variable (composed by the first name type variable and the movement variable) to keep only the movement variable toward North African first names in the analysis. Along with this fixed effect of movement, we computed the random intercepts and related slopes for participants, stimuli, and their interaction. Another model tested whether variability of approach/avoidance tendencies toward French first names was linked to self-reported prejudice—we believed it should not be—by using the same mixed model as for North African first names. In each of these models, we tested whether the random slope of approach/avoidance tendencies (i.e., variability of approach/avoidance tendencies toward either French or North African first names) was linked with prejudice in a classic linear regression model.

Concerning the VAAST, our analysis revealed that self-reported prejudice was significantly linked to variability of approach/avoidance tendencies toward North African first names, $F(1, 44) = 4.17$, $p = .047$. This link was not significant for French first names, $F(1, 44) = 0.70$, $p = .40$. Concerning the Manikin task, the self-reported prejudice was neither linked to variability of approach/avoidance tendencies for North African first names, $F(1, 44) = 0.86$, $p = .36$, nor for French first names, $F(1, 44) = 1.67$, $p = .20$.

2.3 | Discussion

In line with other experiments, in the VAAST, participants were faster in approaching ingroup stimuli (here, French) and avoiding outgroup stimuli (here, North African) rather than the reverse. As predicted, this effect was significantly variable among participants (in both the VAAST and Manikin task), all of this in a population exposed in a similar way to a given cultural knowledge. Interestingly, in the VAAST, variability of approach/avoidance tendencies toward North African first names—but not toward French first names—was linked to self-reported prejudice. In the Manikin task, we did not observe such a link. Because of this pattern of results, we chose to focus on the VAAST to assess intergroup differences in Experiment 2.

3 | EXPERIMENT 2

In this experiment, we tested whether approach/avoidance tendencies differed according to self-reported ethnic origin. Classic indirect attitude tasks (e.g., IAT) are rarely able to capture a possible ingroup bias as a function of group membership, especially when

shared cultural value is associated with the social groups of interest (Ashburn-Nardo et al., 2003; Nosek et al., 2002; Richeson et al., 2005). On the contrary, in the VAAST, given our previous reasoning on approach/avoidance tendencies and the importance of self-reference, we should observe a moderation of the intergroup compatibility effect by participants' ethnic group membership. Precisely, we expected an ingroup bias for both participants of French origin—replicating the results of Experiment 1—and of North African origin, this in spite of similar cultural knowledge about individuals of North African origin (i.e., a negative social value).

3.1 | Method

3.1.1 | Participants and design

To estimate our sample size for sufficient power, we relied on Experiment 1 (N = 50 for a 2 × 2 within-subject by continuous variables). Experiment 2 was also originally designed to test the effect of a between-subjects manipulation[4] on the compatibility effect by participants' self-reported ethnicity interaction. To this end, we estimated a sample size of 150 participants for a design that was ultimately a 2 × 2 mixed design, given that the between-subjects manipulation did not yield any significant effect. Finally, 156 participants took part in this experiment (M_{age} = 20.56, SD_{age} = 2.17, 78 female participants, 42 self-reported North African participants and 22 self-reported ethnicities other than French or North African) in exchange for 10 euros. As in Experiment 1, participants came from the same university. In this experiment, we used a 2 (compatibility: compatible vs. incompatible) × 2 (self-reported origin: French vs. North African) × 2 (block order: compatible first vs. incompatible first) design with the last two variables being between participants.

The design was the same as in the VAAST in Experiment 1. Please note that we kept the same "compatible" and "incompatible" labels as in Experiment 1. What we call compatible (i.e., approaching French first names and avoiding North African first names) and incompatible (i.e., approaching North African first names and avoiding French first names) blocks are compatible and incompatible, respectively, for self-reported French participants, but should be incompatible and compatible, respectively, for self-reported North African participants. For the latter, lower RTs for the incompatible block compared to the compatible block would thus translate into a pro-ingroup (and not a pro-outgroup) compatibility effect.

3.1.2 | Procedure

The procedure was the same as in Experiment 1 except for a few changes: Participants only had to perform the VAAST (and not the Manikin task) and we used a slightly different version of the VAAST. As in some versions of the VAAST used in Rougier et al. (2018, Experiment 4), we removed the character presented in the previous virtual environment, and the visual environment was not a corridor but a 3D street generated in Blender© (for more information see Rougier et al., 2018). As in Experiment 1, we asked participants to indicate whether they considered themselves to be of French, North African, or another origin. Based on this question, we removed individuals answering "other" (N = 22), leaving a total sample of N = 134. We also used the two French-language proficiency questions and all participants reported having a high proficiency level.

3.2 | Results

In this experiment, we predicted an interaction between compatibility and self-reported origin, so that both social groups should present pro-ingroup compatibility effects. As in Experiment 1, we selected RT filters and transformations as a function of the normality of RTs distribution—see the Tables S1–S5 for more information. We removed incorrect trials (i.e., 2.98% of the trials), as well as RTs faster than 350 ms and exceeding 1,600 ms (i.e., 3.49% of the trials). We also removed one participant having more than 96% of errors in the compatible block. Given that the block order did not moderate the compatibility by self-reported origin interaction, $F(1, 128.15) = 2.02$, $p = .16$, but that it did moderate the compatibility effect, $F(1, 128.15) = 16.10$, $p < .001$, we kept this control factor in all our analyses.[5]

The compatibility effect was, on average, non-significant, $F(1, 106.58) = 0.13$, $p = .72$, $dz = 0.09$, and, critically, it was significantly moderated by the self-reported origin of participants, $F(1, 127.84) = 14.31$, $p < .001$, $d = 0.66$ (see Figure 2). A simple effects analysis revealed that self-reported French participants were faster in the compatible block (M = 641 ms, SE = 8 ms) than in the incompatible block (M = 660 ms, SE = 9 ms), $F(1, 108.47) = 7.32$, $p = .008$, $dz = 0.32$. Conversely, self-reported North African participants were faster in the incompatible block (M = 654 ms, SE = 16 ms) compared to the compatible block (M = 667 ms, SE = 16 ms), $F(1, 125.84) = 6.19$, $p = .014$, $dz = 0.36$. Interestingly, these simple compatibility effects varied significantly among both self-reported French, $\chi^2 = 93.01$, $p < .001$ and North African participants, $\chi^2 = 62.43$, $p < .001$.

[4]This between-subjects manipulation was either a "terrorist attack" priming or a control priming (i.e., an airplane crash). We showed to participants a series of portraits describing individuals who supposedly died in the terrorist attack that took place in Paris in 2015 or who supposedly died in an airplane crash (flight 9525 of Germanwings in 2015). We expected the terrorist attack priming to increase the compatibility effect among French participants. Given that this manipulation did not significantly moderate the compatibility effect, $F(1, 124.06) = 0.28$, $p = .59$, or the compatibility by origin interaction, $F(1, 124.06) = 1.23$, $p = .27$, we removed this factor from our analysis. The hypothesis regarding the effect of group membership was formulated in addition to the priming manipulation at the beginning of data collection when noticing the unusually large proportion of North African individuals in the pool of (paid) participants.

[5]Removing this factor did not impact the significance of reported results.

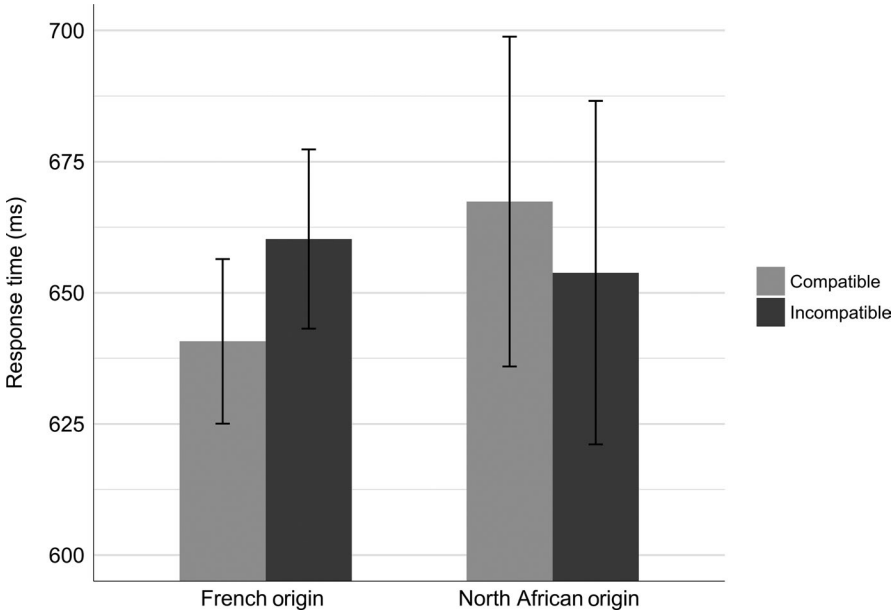

FIGURE 2 Response time (ms) as a function of self-reported origin (French vs. North African) and block (compatible vs. incompatible). Error bars represent the 95% confidence intervals

3.3 | Discussion

Contrary to what is often observed on classic indirect measures—that is, a bias in favor of the dominant (positively valued) social group independently of individuals' group membership—we were able to produce an ingroup bias among stigmatized individuals by using an approach/avoidance task (here, the VAAST). We produced this bias in a homogeneous population (i.e., students of the same university), thus exposed to similar cultural knowledge about social groups. We also replicated the results of Experiment 1 regarding the compatibility effect among individuals of French origin and the associated variability of this effect. Interestingly, the compatibility effect was also variable among North African origin individuals. Together, results of Experiments 1 and 2 are in line with the idea that approach/avoidance tendencies (measured here by the VAAST) may be less influenced by shared cultural knowledge about social groups conveyed in society, as compared to classic indirect measures of attitudes. In a last experiment, we tested whether the compatibility effect in the VAAST could predict scores obtained from ingroup/outgroup trustworthiness ratings. Additionally, we investigated potential differences with the classic IAT measure in predicting these trustworthiness ratings.

4 | EXPERIMENT 3

In Experiment 1, we showed that the inter-individual variability in approach/avoidance tendencies toward North Africans was to some extent explained by self-reported prejudice toward North Africans. Using a direct measure as a criterion variable to validate an indirect measure, however, can be theoretically debated (e.g., regarding dual-process models; Gawronski & Bodenhausen, 2006). Specifically, this finding would not be informative regarding whether approach/avoidance tendencies can correlate with a criterion variable known to share a unique part of variance with indirect measures of prejudice (Stanley et al., 2011). If approach/avoidance tendencies are a valuable assessment of prejudiced attitudes, as we predict, this should be the case. A previous study showed that the IAT can predict discrepancy scores of ingroup/outgroup trustworthiness ratings (Stanley et al., 2011). Accordingly, in Experiment 3, our main objective was to test whether approach/avoidance tendencies can also predict such a discrepancy in trustworthiness ratings. We also explored whether this link goes beyond two self-report measures of prejudice, namely the feeling thermometer and the External Motivation to control prejudice Scale (EMS; Plant & Devine, 1998). This hypothesis was tested in an online setting, allowing a broader sample and the generalization of our effects to different social groups (i.e., African Americans/European Americans). Finally, we also measured IAT scores toward the same social groups and thus investigated potential differences between approach/avoidance tendencies and this classic indirect measure of attitudes in predicting differences in trustworthiness ratings between European American and African American faces.

4.1 | Method

4.1.1 | Participants and design

On the basis of a sensitivity analysis, we aimed for 400 participants because this sample size would be enough to detect a small to medium effect size (f = 0.14), with a power level of 0.80 and an alpha level of 0.05. Accordingly, we recruited 422 participants (M_{age} = 36.42, SD_{age} = 12.58, 226 female participants) in case of exclusion. Participants were recruited via the Prolific Academic platform and were paid 2.14 euros. As pre-selection rules (Prolific Academic filters), we only authorized participants self-defined as white individuals, living in the US, and speaking English as a first language.

Both the VAAST and the IAT follow a 2 (compatibility: compatible vs. incompatible) × 2 (block order: compatible first vs. incompatible

first) × 2 (task order: VAAST first vs. IAT first) design with the last two variables being between participants. The design of the VAAST was the same as in Experiments 1 and 2 except that participants had to approach/avoid European American/African American first names (instead of French/North African first names) and that we used 30 first names (instead of 40). Accordingly, each block comprised 60 trials (plus two practice blocks of 8 trials each over 4 first names not presented elsewhere). The design of the IAT comprised five blocks with three practice blocks (Blocks 1, 2, and 4 of 30 trials each) and two experimental blocks (Blocks 3 and 5 of 60 trials each). Practice blocks consisted of a categorization of positive/negative words (Block 1) or European American/African American first names (Blocks 2 and 4) separately. Experimental blocks consisted of a categorization of both categories of stimuli. We used the same first names (30) as in the VAAST in addition to 30 (15 positive and 15 negative) words, each stimulus being presented once per block. In one experimental block (i.e., compatible), positive words and European African first names shared the same response key (as well as negative words and African American first names), with the opposite configuration for the other block (i.e., incompatible). The order of the two experimental blocks in the IAT was randomized across participants, as well as the task order. The European American/African American first names used in both tasks and the positive/negative words used in the IAT were selected from the original pool of stimuli (100 stimuli) used in Greenwald et al. (1998).

Because in this experiment we tested whether the VAAST and the IAT could predict trustworthiness ratings, we used the score of each task as continuous predictors. Accordingly, we used a continuous (standardized compatibility score in the VAAST: from 0 to 1) × continuous (standardized *D* score in the IAT: from 0 to 1) × continuous (feeling thermometer score: from 1 to 9) × continuous (EMS score: from 1 to 9) × 2 (block order in the VAAST: compatible first vs. incompatible first) × 2 (block order in the IAT: compatible first vs. incompatible first) × 2 (task order: VAAST first vs. IAT first) design with all variables being between participants.

4.1.2 | Procedure

The overall experiment was presented as a "study on person perception and categorization". Participants were informed that they would have to perform several tasks on face perception and first name categorization, followed by a questionnaire at the end of the experiment. Participants first completed the trustworthiness rating task, after which they had to perform the VAAST and the IAT (in a randomized order). Finally, they filled in the feeling thermometer and the EMS questionnaire (in a randomized order), and answered demographic questions.

4.1.3 | Trustworthiness ratings

The procedure was similar to the one used by Stanley et al. (2011) who showed a correlation between discrepancies in trustworthiness ratings (toward ingroup/outgroup faces) and IAT scores (toward European American/African American individuals). Participants rated 60 male faces selected from the Chicago database (Ma, Correll, & Wittenbrink, 2015). Among these 60 faces, we selected 20 European American faces, 20 African American faces, 10 Asian American faces, and 10 Latino American faces. We selected European American and African American faces so that the two groups did not differ significantly on their level of trustworthiness, $t(58) = 0.04$, $p = .97$, and prototypicality, $t(58) = 0.01$, $p = .99$. Faces were displayed one by one during 1 s and, for each trial, participants had to rate the face on a Likert scale from 1 (*not-at-all trustworthy*) to 9 (*extremely trustworthy*), as in Stanley et al. (2011). We only analyzed ratings for European American and African American faces (cf. Stanley et al., 2011).

4.1.4 | VAAST

The procedure was similar to the one of Experiment 2, except for two differences. First, we used a different version of the VAAST (the same as in Experiment 4 of Rougier et al., 2018). In this version, participants only had to press once the approach/avoidance key to categorize stimuli, simulating visually a short movement forward/backward in the environment. This one-key-press version of the VAAST has the advantage of diminishing the overall time of completion and, crucially, it is still able to produce large approach/avoidance effects in lab experiments (Rougier et al., 2018) as well as in online experiments (Aubé, Rougier, Muller, Ric, & Yzerbyt, 2019). Second, because the experiment was administered online and not in the lab, participants categorized the first names with their keyboard (instead of a button box). Specifically, participants used the H key as a start button, the Y key to move forward, and the N key to move backward.

4.1.5 | IAT

We used the 5-block version of the IAT (as in Greenwald et al., 1998). In this task, participants were asked to categorize stimuli using the E and I keys of their keyboard, as fast as they could. In case of error, a red cross was displayed and participants were asked to correct their response with the opposite key. The practice blocks (Blocks 1, 2, and 4) consisted of the categorization of positive/negative words or European American/African American first names. As in the classic IAT procedure, valenced categories were always associated with the same response key (e.g., positive words were always associated with the E key), but first name categories association with response keys switched from Block 2 to Block 4. Key assignment was randomized across participants. The experimental blocks (Blocks 3 and 5) combined all stimuli and could either be compatible (e.g., European American first names and positive words sharing the same key) or incompatible (e.g., European American first names and positive words sharing the opposite key). We recorded the response time for each trial (time between stimuli display and participants' response), but only analyzed those of experimental blocks.

4.1.6 | Feeling thermometer and EMS scales

In the feeling thermometer scale, participants indicated "how warm or cold" they felt toward black people and white people (in a randomized order), from 1 (*coldest feelings*) to 9 (*warmest feelings*). In the EMS scale, participants indicated to what extent they agreed, from 1 (*strongly disagree*) to 9 (*strongly agree*), with 5 statements (in a randomized order) measuring the external motivation to control prejudice (e.g., "I attempt to appear non-prejudiced toward Black people in order to avoid disapproval from others").

Finally, participants answered demographic questions (sex and gender). We also asked them to report their racial/ethnic category (across 5 options: American Indian or Alaska Native; Asian, Black or African American, Hispanic or Latino, Native Hawaiian or Other Pacific Islander, and White) and their English proficiency. All participants reported having a high proficiency level.

4.2 | Results

First, our main prediction was that the compatibility effect in the VAAST should correlate with trustworthiness ratings. In other words, we predicted that an increase on the difference score between incompatible block and compatible block in the VAAST (i.e., compatibility effect = RT incompatible − RT compatible) would predict an increase in the difference in trustworthiness ratings toward European American versus African American people (i.e., trustworthiness score = trust European American − trust African American). We also explored whether this link could go beyond direct measures of prejudice (i.e., whether this correlation was still significant by adding the feeling thermometer and the EMS measures in two separate models). Second, we investigated whether the link with trustworthiness ratings was larger for the VAAST compatibility score as compared to the IAT *D* score.[6] Performing a mixed model for these analyses was not possible given that we used aggregated scores based on RT to predict trustworthiness ratings (instead of predicting RT with the VAAST compatibility variable as in Experiments 1 and 2). Accordingly, we relied on OLS regression analyses. In the main analysis, we considered the compatibility score in the VAAST as a between-participants variable. Conversely, in the secondary analysis—i.e., when comparing the VAAST and the IAT in predicting trustworthiness ratings—we considered VAAST versus IAT scores as a within-participant variable comparing the predictive power of the two tasks.

Finally, although this was not the aim of this experiment, we investigated the variability of the VAAST compatibility effect and the one of the IAT *D* score across participants in two separate mixed model analyses—note that we did not pre-register such an analysis. For the VAAST, we followed the same mixed model as in Experiments 1 and 2. For the IAT, we followed Wolsiefer et al's (2016) recommendations. Specifically, we estimated a mixed model having the compatibility of the block (compatible vs. incompatible), the type of category (valenced words vs. social groups), the type of word (positive vs. negative), and the type of group (African American vs. European American) as fixed effects and we estimated the random intercepts and the compatibility slope for participants and stimuli.

On the basis of our pre-registered exclusion criteria, we excluded from all analyses 14 participants having more than 40% of errors in the VAAST (i.e., 3.32% of the sample), 11 participants for whom more than 10% of their trials had a latency inferior to 300 ms in the IAT (i.e., 2.69% of the sample; following Greenwald, Nosek, & Banaji, 2003), and 3 participants who reported being non-white (recruited by mistake), leaving us with 394 participants. As in Experiments 1 and 2, we selected RT filters and transformations for the VAAST as a function of the normality of RTs distribution—see the Tables S1–S5 for more information. Accordingly, we removed incorrect trials (i.e., 4.32% of the trials), as well as RTs faster than 450 ms and exceeding 1,800 ms (i.e., 3.59% of the trials). Regarding IAT data, following Greenwald et al. (2003) recommendations, we excluded trials with latencies superior to 10,000 ms (i.e., 0.04% of the trials) and we replaced each error latency by the block mean added to 600 ms (i.e., 8.13% of the trials).

4.2.1 | VAAST in predicting trustworthiness ratings

We used the compatibility scores in the VAAST (i.e., compatibility score = RT incompatible − RT compatible) as a predictor of trustworthiness scores (i.e., European American faces ratings − African American faces ratings). Neither block order, $F(1, 384) = 0.89$, $p = .35$, nor task order, $F(1, 384) = 0.63$, $p = .43$, had a significant impact on our results.[7] Accordingly, we removed these control factors from the following analyses.

First, the compatibility effect obtained in Experiments 1 and 2 with French participants was replicated here with an American sample and different target groups: Participants were significantly faster in the compatible block ($M = 774$ ms, $SE = 7$ ms) than in the incompatible block ($M = 788$ ms, $SE = 7$ ms), $F(1, 387) = 8.10$, $p = .005$, $dz = 0.14$. This compatibility effect varied significantly across participants, $\chi^2 = 942.32$, $p < .001$, as attested by the separate mixed model analysis. Second, and more important, the compatibility effect significantly predicted discrepancies in trustworthiness ratings, $b = 0.09$, $F(1, 386) = 3.98$, $p = .047$, $\eta^2_p = 0.010$, indicating that the larger the compatibility effect (i.e., the larger the bias in favor of European Americans), the larger the discrepancy in trustworthiness ratings (i.e., in favor of European Americans). Finally, this link remained significant when adding EMS scores in the model, $b = 0.09$, $F(1, 385) = 4.11$, $p = .043$, $\eta^2_p = 0.011$. When adding the feeling thermometer scores (i.e., thermometer score = warmth feelings toward white

[6]One stimulus in the IAT procedure was miscoded (i.e., the first name "Harry" was coded as an African American first name instead of a European American first name). Accordingly, we removed this stimulus for all analyses using the IAT data.

[7]For all the analyses of this section, we removed 5 outliers (having |SDR| > 4) from the data. Removing these outliers did not impact the significance of the presented results.

people – warmth feeling toward black people), however, this link became marginal, $b = 0.07$, $F(1, 385) = 2.87$, $p = .091$, $\eta^2_p = 0.007$.[8,9]

4.2.2 | VAAST versus IAT comparison

To compare the VAAST and the IAT in predicting trustworthiness ratings, we computed a new score, being the difference between the standardized compatibility score in the VAAST and the standardized *D* score in the IAT (difference score = standardized VAAST compatibility score – standardized IAT *D* score). We used this difference score as a predictor of the discrepancy of trustworthiness scores (i.e., European American faces ratings – African American faces ratings). Accordingly, our model tested to what extent the VAAST-IAT difference predicted discrepancies of trustworthiness rating. If, for instance, the VAAST predicted discrepancies of trustworthiness ratings to a larger extent than the IAT, then we would find a positive regression parameter, conceptually equivalent to a type of indirect measure by the strength of the bias interaction. Neither block order in the VAAST, $F(1, 384) = 0.42$, $p = .52$, block order in the IAT, $F(1, 384) = 0.02$, $p = .90$, nor task order, $F(1, 384) = 0.65$, $p = .42$, had a significant effect.[10] We thus removed these control factors from the following analyses.

Our analysis showed that the difference between the compatibility score in the VAAST and the *D* score in the IAT did not significantly impact trustworthiness scores, $b = -0.03$, $F(1, 386) = 0.86$, $p = .35$, $\eta^2_p = .002$. In other words, although descriptively in favor of the IAT *D* scores (the regression parameter being negative), the difference in predictive power was not significant. Further analyses showed that the average *D* score was moderately in favor of European Americans ($M = 0.54$, $SE = 0.02$), $F(1, 387) = 659.49$, $p < .001$, $dz = 1.30$. This compatibility effect varied significantly across participants, $\chi^2 = 11{,}792$, $p < .001$, as attested by the separate mixed model analysis. In line with Stanley et al. (2011), the IAT *D* scores significantly predicted trustworthiness ratings, $b = 0.14$, $F(1, 386) = 10.12$, $p = .002$, $\eta^2_p = 0.026$, so that the larger the *D* score, the larger the rating difference between European American and African American faces. This effect remained significant when controlling for EMS, $b = 0.14$, $F(1, 385) = 9.97$, $p = .002$, $\eta^2_p = 0.025$, or the feeling thermometer, $b = 0.09$, $F(1, 385) = 5.86$, $p = .02$, $\eta^2_p = 0.015$.

4.3 | Discussion

To the best of our knowledge, this experiment is the first to test the relationship between approach/avoidance tendencies and trustworthiness evaluations of ingroup/outgroup members. As predicted, we showed that the more biased the compatibility effect of approach/avoidance, the more biased the trustworthiness ratings of ingroup compared to outgroup members. Significantly, this link remained quite consistent when self-report measures were added in the analysis—with more mixed results for the feeling thermometer, but see the Section 5 for further discussion. These results thus corroborate and extend Experiment 1's findings, showing that approach/avoidance tendencies toward social groups are linked to a variable known to share a unique relationship with an indirect measure of prejudice. In addition, and also for the first time, we investigated potential differences between approach/avoidance tendencies and the IAT *D* score in predicting a criterion variable like trustworthiness ratings. Contrary to what we would have expected, however, we did not observe a significant difference in terms of predictive power between these two measures. Implications of this last finding are discussed in the Section 5. Finally, in two separate mixed model analyses, both VAAST compatibility effect and IAT *D* score varied significantly across participants. Although these analyses were not the focus of the present experiment and the relevance of such analyses (given our non-homogeneous population) can be discussed, it should be noted that this is the first time that participants' variability on the IAT *D* score is tested directly. Of course, studies examining the predictive validity of the IAT at the individual level (e.g., by correlating the *D* score with direct measures; Hofmann et al., 2005, or behaviors; Kurdi et al., 2019) already attest, to some extent, to such inter-individual variability. Yet, a mixed model analysis has the advantage of both quantifying this variability and not requiring an additional measure (therefore avoiding the risk that comes with adding a variable to assess this variability).

5 | GENERAL DISCUSSION

In this work, we addressed the question of whether approach/avoidance tasks are affected by an element of criticism often addressed at classic indirect measures of attitudes, namely whether they could capture personal attitudes beyond cultural knowledge (Arkes & Tetlock, 2004; Karpinski & Hilton, 2001; Payne et al., 2017). We explored whether approach/avoidance tasks could be an interesting alternative regarding this issue. If they are, we reasoned that they should produce two kinds of results: Individual-level variability and group-level variability. Specifically, approach/avoidance tasks should be able to discriminate individuals on their approach/avoidance scores—that is, to produce approach/avoidance compatibility effects varying across individuals—and these tasks should be able to produce approach/avoidance effects depending on individuals' actual social group membership—that is, to produce pro-ingroup compatibility effects even for groups associated with (negatively-valenced) cultural knowledge. We addressed these two points in two experiments (Experiments 1 and 2) and we did so among a population

[8] In each of these analyses, EMS scores, $F(1, 385) = 10.05$, $p = .002$, $\eta^2_p = .025$, and feeling thermometer scores, $F(1, 385) = 129.85$, $p < .001$, $\eta^2_p = .252$, significantly predicted trustworthiness ratings. These links were also significant when testing without the compatibility effect for both EMS scores, $F(1, 386) = 9.95$, $p = .002$, $\eta^2_p = .025$, and feeling thermometer scores, $F(1, 386) = 161.44$, $p < .001$, $\eta^2_p = .293$.

[9] Because this could be of interest for researchers, we tested whether EMS scores moderate the relationship between the racial biases in the VAAST and the IAT and trustworthiness ratings (in two separate analyses). Results showed that EMS scores significantly moderate the relationship between the VAAST compatibility effects and trustworthiness ratings (so that the higher the EMS scores, the lower the relationship between compatibility effects and trustworthiness ratings), $F(1, 384) = 4.35$, $p = .04$, $\eta^2_p = .011$, but not the one between IAT *D* scores and trustworthiness ratings, $F(1, 384) = 0.84$, $p = .77$, $\eta^2_p < .001$.

[10] For all the analyses of this section, we removed 5 outliers (having |SDR| > 4; different from the previous analysis section) from the data. Removing these outliers did not impact the significance of presented results.

likely to be exposed to a similar cultural knowledge about French and North African groups (i.e., students at the same university). In a last experiment (Experiment 3), we tested whether approach/avoidance tendencies could correlate with a criterion variable (i.e., trustworthiness ratings) known to share a unique part of variance with an indirect measure of prejudice. We also contrasted the predictive power of approach/avoidance tendencies measured with the VAAST with that of the classic intergroup bias measured with the IAT.

Experiment 1 addressed the question of the individual-level variability among participants of (self-reported) French origin. We tested whether two approach/avoidance tasks known to produce large compatibility effects (i.e., the VAAST and the Manikin task) were able to produce an intergroup compatibility effect (i.e., faster response time to approach French first names and to avoid North African first names than the reverse) that was variable across individuals. In line with other intergroup approach/avoidance experiments (e.g., Paladino & Castelli, 2008), we produced a significant compatibility effect with the VAAST. In the Manikin task, this effect was not significant (see Rougier et al., 2018 for a similar pattern). Crucially and as predicted, this effect was variable across individuals with some participants having a stronger/weaker ingroup bias. In addition, for the VAAST, variability of approach/avoidance actions toward North African stimuli (but not toward French stimuli) was linked to participants' self-reported prejudice toward North Africans, attesting to the meaningfulness of this variability (see Fazio et al., 1995, for a similar reasoning).

Experiment 2 addressed the question of the group-level variability, that is, whether individuals belonging to dominant versus non-dominant groups can both produce a compatibility effect in favor of their respective group. To do so, we recruited participants categorizing themselves as being of French versus of North African origin. Using the VAAST, both groups of participants produced, on average, a compatibility effect in which they were faster in approaching ingroup stimuli and in avoiding outgroup stimuli rather than the reverse. Furthermore, we observed significant variability of the compatibility effect across both French (replicating Experiment 1) and North African groups of participants.

In Experiment 3, we showed that the compatibility effect produced by the VAAST could also correlate with a subtler measure of prejudice usually correlated with indirect tasks beyond direct assessment of prejudice. Specifically, the more biased the compatibility effect of approach/avoidance was, the more biased were the trustworthiness ratings toward the ingroup relative to the outgroup. This effect remained independent of self-report measures of prejudice. Contrary to what we expected, however, we did not find significant differences between the VAAST and the IAT in predicting trustworthiness ratings.

5.1 | Contribution to indirect measures of intergroup attitudes

These results are of particular interest regarding the value of the VAAST in measuring individuals' intergroup attitude. First, in all our experiments, the intergroup compatibility effect was significant in the VAAST. Although this effect has been found several times in the literature, it was often found with tasks relying on arm movements (e.g., Clow & Olson, 2010; Paladino & Castelli, 2008; Vaes et al., 2003). However, arm movements are ambiguous to interpret because arm flexion, for instance, can be used both to bring a stimulus toward us and to move our hand away from the stimulus (the same being true for arm extensions; Seibt et al., 2008). And indeed, using a similar setting with a modified keyboard, Paladino and Castelli (2008) interpreted arm flexions and arm extensions respectively as avoidance and approach, while Alexopoulos and Ric (2007) interpreted (along with Chen & Bargh, 1999) arm flexion and arm extension respectively as approach and avoidance. Because tasks like the VAAST and the Manikin task simulate movements of the whole self these tasks do not raise such interpretative issues (see Rougier et al., 2018). Replicating the ingroup/outgroup compatibility effect with such a task (here the VAAST) is therefore a first contribution.

In the area of intergroup attitudes, however, the ability of a task to produce intergroup (compatibility) effects does not equate with being a good measure of intergroup attitudes. As discussed in the next section, a widespread and large effect could rather be explained by a situational (e.g., institutional) bias than by a personal one (Arkes & Tetlock, 2004; Payne et al., 2017). Accordingly, a compatibility effect that is variable across individuals, as observed in Experiments 1 and 2, is a more informative result regarding the relevance of a task as a measure of attitudes. This result is informative because cultural knowledge is likely to be invariant (see Fazio et al., 1995 for a similar reasoning)—especially in a homogeneous student population. In the same vein, both ingroup biases in the VAAST and the IAT also varied across individuals in Experiment 3; yet these results have to be interpreted with caution because the population was not homogeneous in this experiment—this variability could thus be interpreted as resulting from variability in cultural knowledge.

Although the variability is informative, it is not sufficient by itself (except in the case of the absence of variability): This one has to be meaningfully explained by a relevant individual-level variable, as we did in the present contribution. One could argue, indeed, that the impact of variables other than personal prejudice could still produce a certain level of variability. First, some variability in exposure to cultural information could still persist across individuals because their immediate environment cannot be exactly the same (even for participants like ours from the same university, that is, individuals living in a similar immediate environment). For instance, individuals could vary on their momentary exposure to negative versus positive information about social groups (e.g., somebody in my environment just said that North Africans are aggressive) or on their chronic exposure to this information (e.g., my family relays negative information about North Africans; see Payne et al., 2017 for a similar line of reasoning). Variability on these two kinds of environmental (momentary or chronic) exposure could possibly account for variability of approach/avoidance effects. Second, some variability on other individual-level variables, such as cognitive abilities to perform the approach/avoidance task (e.g., executive functions), could also have contributed to

this inter-individual variability. Yet, we observed that variability of approach/avoidance tendencies toward North African first names was meaningfully linked to a relevant individual-level variable: Self-reported prejudice toward North African individuals (Experiment 1). This result is thus consistent with the idea that variability of approach/avoidance tendencies (here, toward North Africans) reflected variability of personal attitudes rather than variability in terms of (momentary or chronic) exposure to knowledge about social groups. Additionally, Experiment 3 showed that the compatibility effect in the VAAST can also correlate with a subtler measure of prejudice known for its unique relationship with an indirect measure, namely trustworthiness evaluations of ingroup/outgroup members.

These results, however, have to be interpreted with caution given the p-values obtained when assessing the link between approach/avoidance tendencies and individual-level variables (i.e., self-report prejudice in Experiment 1 and trustworthiness evaluations in Experiment 3). Indeed, even if one of these two experiments was pre-registered (i.e., Experiment 3), the p-values were very close to the critical threshold of .05. We must therefore direct readers' attention to these p-values, which, according to some authors, should be interpreted with care (e.g., Benjamin et al., 2018). Accordingly, we believe that future work should corroborate these results to ascertain their robustness.

Second, the VAAST was able to produce a pro-ingroup compatibility effect among individuals from both dominant and non-dominant groups belonging to a homogeneous population of university students. This result contrasts with some IAT results in which people display a biased effect in favor of the dominant group (e.g., in favor of white individuals; Nosek et al., 2002). In the case of our present focus on approach/avoidance tasks as indirect attitude measures, this result is to be linked with two recent experiments published after we conducted our research. A first experiment compared students from segregated versus non-segregated schools by using the Manikin task (Degner et al., 2016). In this experiment, minority students from segregated schools (i.e., from schools with a majority of students from immigration background) displayed a pro-ingroup compatibility effect, whereas minority students from non-segregated schools (i.e., from schools with a majority of white students) displayed a pro-outgroup compatibility effect. This last result (i.e., a pro-outgroup compatibility effect among minority students from diverse schools) contrasts with what we obtained in Experiment 2, that is, a pro-ingroup compatibility effect among individuals of North African origin studying in a diverse environment. The second experiment compared White (dominant) and Black (non-dominant) individuals in a Portuguese context (i.e., a diverse context) by using the modified keyboard task (Bianchi et al., 2018). In line with our results, these authors showed that both groups produced a pro-ingroup compatibility effect.

Several factors could account for the difference between Degner et al.'s results (i.e., a pro-outgroup compatibility effect) on one side and Bianchi et al.'s and ours (i.e., an ingroup bias in Experiment 2) on the other side. First, Bianchi et al.'s population and ours were comprised of adults (e.g., university students in our experiment) and not adolescents. Adolescence is a crucial period for conformity development and adoption of peers' behaviors (e.g., Costanzo & Shaw, 1966; Krosnick & Judd, 1982). Accordingly, a possibility is that developmental stage (here, adolescence) impacted the way individuals (notably individuals belonging to non-dominant groups) conform to the behaviors of the dominant group, among them approach/avoidance behaviors (impacting approach/avoidance tendencies measured by the task). Second, Bianchi et al. and we relied on a task involving sensorimotor aspects of real-world approach/avoidance behaviors (e.g., arm movements of flexion/extension for Bianchi et al., 2018, and visual aspects of the whole self moving forward/backward in the present work), which is not the case for the Manikin task (relying on more symbolic approach/avoidance actions). For now, it is unclear what factors could intervene in making approach/avoidance tasks more appropriate to capture individuals' attitude beyond cultural knowledge. Therefore, we cannot exclude that the Manikin task is more sensitive than the VAAST to cultural knowledge. One critical aspect could be whether the task implements sensorimotor aspects reproducing real-world approach/avoidance behaviors (e.g., Rougier et al., 2018). Even if, for now, we cannot explain the difference between Degner et al.'s results on one side and Bianchi et al.'s and ours on the other, it is important to note that the VAAST is able to capture a pro-ingroup compatibility effect among non-dominant individuals belonging to the same (diverse) environment as dominant individuals. Future work should investigate whether this compatibility effect can also be found among adolescents and compare the VAAST and the Manikin task in doing so.

5.2 | Should we use the VAAST instead of classic indirect tasks? Current debate and future directions

The question of whether one should use the VAAST instead of classic indirect tasks arises from our reasoning and from the empirical evidence provided by Experiments 1 and 2. Even if this question was not the focus of the present article, we conducted Experiment 3 to investigate this possibility and to provide some preliminary answers. This experiment aimed at contrasting the VAAST with a classic indirect task, that is, the IAT, in predicting a subtle measure of prejudice. Extending the results of Experiments 1 and 2, we found that the approach/avoidance bias measured by the VAAST correlated with a criterion variable known to share a unique part of variance with an indirect measure. Contrary to what we would have expected, however, this effect was not significantly different from the one obtained with the IAT. These results thus nuance those obtained in Experiments 1 and 2 regarding the special status of the VAAST as compared to classic indirect tasks.

On the basis of these last results comparing the VAAST and the IAT, it could be argued that approach/avoidance tasks do not capture personal prejudice to a larger extent than classic indirect measures do. Our reasoning was that, because approach/avoidance tasks assess embodied reactions (and not positive/negative associations; see Experiment 5 in Rougier et al., 2018), approach/avoidance tendencies should be a better predictor of

personal prejudice, as measured for instance by trustworthiness ratings. Experiment 3's results, however, do not support this view. Accordingly, it could be argued that approach/avoidance tasks like the VAAST only capture the positive/negative dimension associated with social group as the IAT does, and not embodied—i.e., more personal—approach/avoidance reactions. This interpretation is consistent with some empirical work showing that approach/avoidance implementation is largely confounded with a positive/negative valence, respectively (e.g., Eder & Klauer, 2009; Eder, Rothermund, De Houwer, & Hommel, 2015; Eder & Rothermund, 2008; but see Rougier et al., 2018, Experiment 5).

Whereas Experiment 3's results do not support the idea of the VAAST's superiority in predicting personal prejudice, we believe these results do not constitute a definitive answer because of the trustworthiness measure we used to contrast the VAAST and the IAT. We chose trustworthiness evaluations as a measure of personal prejudice precisely because this task shares a unique part of variance with the IAT and thus because it allowed for a conservative comparison between the VAAST and the IAT—given that the IAT *D* score has been shown to highly correlate with this kind of measure (Oswald et al., 2013; Stanley et al., 2011), contrary to the VAAST compatibility effect (that has never been correlated with trustworthiness ratings until now). However, the trustworthiness/untrustworthiness dimension is also strongly related to the valence dimension in the domain of face perception (similar to the measure of face ratings we used), which might overlap with cultural knowledge (i.e., trustworthy = positive; Oosterhof & Todorov, 2008; see also Oliveira, Garcia-Marques, Garcia-Marques, & Dotsch, 2019 for the domain of social judgment). Additionally, we know from meta-analyses on IAT effects that its predictive power with measures of intergroup attitudes—as self-report, response time tasks, behavioral measures, or brain activity—is unequal (Greenwald et al., 2009; Hofmann et al., 2005; Kurdi et al., 2019; Oswald et al., 2013). Accordingly, if the VAAST captures something more than the mere valence associated with social groups, as compared to the IAT, the compatibility effect should be a better predictor than the *D* score for some measures than others, depending on their relationship with valence. We thus recommend further investigation with various measures of intergroup attitudes (e.g., strongly associated with valence or not) as well as caution in interpreting the present results.

Another possibility to contrast the VAAST and the IAT in assessing personal attitudes is to compare these tasks in producing racial bias variability in a homogeneous population. As stated before, however, such a test is not truly informative in the present Experiment 3. Because we aimed at maximizing power as well as generalizing VAAST results to another population (and thus intergroup targets), we went for an online setting. This setting, however, does not provide an optimal setting to test differences between the VAAST and the IAT in terms of inter-individual variability, because any differences on inter-individual variability could be explained by differences on cultural knowledge exposure (i.e., participants coming from various US states/cities). As a consequence, one cannot exclude that the variability in the VAAST and the IAT are of a different nature. For instance, it could be that whereas the variability in the VAAST is due to differences in personal attitudes, the variability in the IAT is due to differences in cultural knowledge (or the other way around, but the literature already points to such differences for the IAT). In line with this idea, future research could contrast approach/avoidance tendencies and classic indirect measures on their inter-individual variability in a homogeneous population (e.g., university students) and by using mixed model analyses as we did in Experiments 1. Such an experiment would not suffer from issues potentially raised in using an individual-level measure of prejudice to distinguish the two tasks. Moreover, we believe that such an investigation would contribute to the emerging debate on whether classic indirect measures of attitudes are better suited to capture individual-level or aggregate-level bias (e.g., Payne et al., 2017).

In a recent debate, Payne et al. (2017) strongly questioned the value of classic indirect tasks (as the IAT and the AMP) in measuring individual biases and argued that these tasks would be more valuable in measuring "systemic biases". According to these authors, implicit bias effects could be due to both chronic and situational accessibilities of concepts in memory—where the chronic accessibility refers to the traditional notion of attitude (i.e., a stable construct; Banaji, 2001). Payne et al. (2017) reviewed empirical evidence showing that effects at the individual level are generally unstable (Gawronski, Morrison, Phills, & Galdi, 2017) and only weakly associated with individual differences (e.g., in predicting discriminatory behaviors; Greenwald et al., 2009; Oswald et al., 2013). On the contrary, at the aggregate level (e.g., IAT scores aggregated as a function of US states), these effects are stable (Payne et al., 2017; see also Hehman, Calanchini, Flake, & Leitner, 2019) and strongly associated with situational variables (e.g., city-level; Zerhouni et al., 2016; region-level; Hehman et al., 2018; county/state-level; Orchard & Price, 2017; Payne et al., 2019). Authors concluded that these effects would be mostly due to the situational accessibility of concepts (i.e., conveyed by racist institutions and cultures), what the authors called the "systemic bias". Accordingly, classic indirect measures of biases, such as the IAT, would be more efficient in predicting "crowds" behaviors (e.g., Hehman et al., 2018; Zerhouni et al., 2016) than individual ones.

The question of whether the VAAST, and more generally approach/avoidance tasks, assess individual rather than "crowds" behaviors and/or is distinct from classic indirect tasks remains open. Whereas the results of Experiments 1 and 2 suggest that the VAAST could be a good measure to assess individual-level prejudice, the results of Experiment 3 also suggest that the VAAST would not differ from classic indirect measures. This experiment, however, is the only one comparing an approach/avoidance task with a classic indirect measure. Accordingly, we believe that to decide whether the VAAST—or, more broadly approach/avoidance tasks—should be used instead of classic indirect measures, one could address the issues raised by Payne et al. (2017). Specifically,

two main issues ought to be addressed: (a) whether the compatibility effect produced by the VAAST is systematic at the individual level (e.g., temporal stability; Gawronski et al., 2017), and (b) whether it can predict individual behaviors (e.g., non-verbal behaviors toward an outgroup member; Word, Zanna, & Cooper, 1974) better than the compatibility effect at the aggregate level (e.g., compatibility scores aggregated at the city-level). To test this last idea, both individual (e.g., coming from various cities) and aggregate (e.g., city) predictors could be modeled in the same multi-level statistical model (e.g., Judd et al., 2012; Westfall et al., 2014) in predicting an individual-level behavior. Both of these approaches could further test the idea that the VAAST is a suitable alternative when it comes to measuring prejudiced attitudes.

6 | CONCLUSION

Indirect measures of attitudes abound in the literature. When it comes to attitudes for which a strong cultural value is associated (as intergroup attitudes), however, these tasks are not equal in assessing personal attitudes beyond cultural influences. In this work, we provide evidence that the VAAST was able to discriminate across individuals and social groups beyond the influence of cultural knowledge. Indeed, these results were shown among individuals belonging to both dominant and non-dominant social groups homogeneously exposed to a pro-dominant cultural knowledge. Even if more research is needed to decide whether the VAAST is a useful alternative to classic indirect measures of attitudes, showing that this task can overcome a central criticism often addressed to them opens new avenues for research on normatively sensitive attitudes.

ACKNOWLEDGEMENTS

This research was part of Marine Rougier's thesis under the supervision of Dominique Muller and Annique Smeding. It was supported by a grant from the Agence Nationale de la Recherche (ANR-2010-BLAN-1905-01), by the CNRS, by the Pôle Grenoble Cognition, by the SFR Santé & Société, and this work was also supported by the French National Research Agency in the framework of the "Investissements d'avenir" program (ANR-15-IDEX-02).

CONFLICT OF INTEREST

Authors declare having no conflicts of interest in publishing these data.

ETHICAL APPROVAL

This work was conducted in line with the ethical guidelines specified in the APA Code of Conduct as well as with authors' national ethics guidelines.

TRANSPARENCY STATEMENT

The data and the R scripts for all the reported experiments can be found at http://bit.do/fbgj8.

ORCID

Marine Rougier https://orcid.org/0000-0002-9467-2726
Dominique Muller https://orcid.org/0000-0001-9544-5317
Rémi Courset https://orcid.org/0000-0002-4165-1411
Annique Smeding https://orcid.org/0000-0002-7743-9578
Thierry Devos https://orcid.org/0000-0003-4546-7555
Cédric Batailler https://orcid.org/0000-0003-0553-6827

REFERENCES

Alexopoulos, T., & Ric, F. (2007). The evaluation-behavior link: Direct and beyond valence. *Journal of Experimental Social Psychology*, *43*(1), 1010–1016. http://doi.org/10.1016/j.jesp.2006.10.017

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

Arkes, H. R., & Tetlock, P. E. (2004). Attributions of implicit prejudice, or "Would Jesse Jackson 'fail' the Implicit Association Test?". *Psychological Inquiry*, *15*(4), 257–278. https://doi.org/10.1207/s15327965pli1504_01

Ashburn-Nardo, L., Knowles, M. L., & Monteith, M. J. (2003). Black Americans' implicit racial associations and their implications for intergroup judgment. *Social Cognition*, *21*(1), 61–87. https://doi.org/10.1521/soco.21.1.61.21192

Aubé, B., Rougier, M., Muller, D., Ric, F., & Yzerbyt, V. (2019). The online-VAAST: A short and online tool to measure spontaneous approach and avoidance tendencies. *Acta Psychologica*, *201*, 1–10. https://doi.org/10.1016/j.actpsy.2019.102942

Banaji, M. R. (2001). Implicit attitudes can be measured. In H. L. Roediger, J. S. Nairne, I. Neath & A. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 117–150). Washington, DC: American Psychological Association.

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., … Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour*, *2*(1), 6–10. https://doi.org/10.1038/s41562-017-0189-z

Bianchi, M., Carnaghi, A., & Shamloo, S. E. (2018). Intergroup attitudes accessibility and motor approach-avoidance responses in White and Black individuals in Portugal. *Psicologia Sociale*, *13*(2), 147–164. https://doi.org/10.1482/90780

Cacioppo, J. T., Priester, J. R., & Berntson, G. G. (1993). Rudimentary determinants of attitudes: II. Arm flexion and extension have differential effects on attitudes. *Journal of Personality and Social Psychology*, *65*(1), 5–17. https://doi.org/10.1037/0022-3514.65.1.5

Castelli, L., Zogmaister, C., Smith, E. R., & Arcuri, L. (2004). On the automatic evaluation of social exemplars. *Journal of Personality and Social Psychology*, *86*(3), 373–387. https://doi.org/10.1037/0022-3514.86.3.373

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, *25*(2), 215–224. https://doi.org/10.1177/0146167299025002007

Clow, K. A., & Olson, J. M. (2010). Conceptual-motor compatibility and homonegativity: Approaching and avoiding words associated with homosexuality. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement*, *42*(4), 222–233. https://doi.org/10.1037/a0020010

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology*, *83*(6), 1314–1329. https://doi.org/10.1037/0022-3514.83.6.1314

Costanzo, P. R., & Shaw, M. E. (1966). Conformity as a function of age level. *Child Development*, *37*(4), 967–995. https://doi.org/10.2307/1126618

Cousijn, J., Goudriaan, A. E., & Wiers, R. W. (2011). Reaching out towards cannabis: Approach-bias in heavy cannabis users predicts

changes in cannabis use. *Addiction*, *106*(9), 1667–1674. https://doi.org/10.1111/j.1360-0443.2011.03475.x

Dambrun, M., & Guimond, S. (2001). La théorie de la privation relative et l'hostilité envers les Nord-Africains [The theory of relative deprivation and hostility towards North Africans]. *Revue Internationale De Psychologie Sociale*, *14*(1), 57–89.

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

De Houwer, J., Crombez, G., Baeyens, F., & Hermans, D. (2001). On the generality of the affective Simon effect. *Cognition & Emotion*, *15*(2), 189–206. https://doi.org/10.1080/02699930125883

Degner, J., Essien, I., & Reichardt, R. (2016). Effects of diversity versus segregation on automatic approach and avoidance behavior towards own and other ethnic groups: Diversity and automatic intergroup behavior. *European Journal of Social Psychology*, *46*(6), 783–791. https://doi.org/10.1002/ejsp.2234

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

Eder, A. B., & Klauer, K. C. (2009). A common-coding account of the bidirectional evaluation–behavior link. *Journal of Experimental Psychology: General*, *138*(2), 218–235. https://doi.org/10.1037/a0015220.

Eder, A. B., & Rothermund, K. (2008). When do motor behaviors (mis) match affective stimuli? An evaluative coding view of approach and avoidance reactions. *Journal of Experimental Psychology: General*, *137*(2), 262–281. https://doi.org/10.1037/0096-3445.137.2.262

Eder, A. B., Rothermund, K., De Houwer, J., & Hommel, B. (2015). Directive and incentive functions of affective action consequences: an ideomotor approach. *Psychological Research Psychologische Forschung*, *79*(4), 630–649. https://doi.org/10.1007/s00426-014-0590-4

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

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

Gawronski, B., & Bodenhausen, G. V. (2007). Unraveling the processes underlying evaluation: Attitudes from the perspective of the APE model. *Social Cognition*, *25*(5), 687–717. https://doi.org/10.1521/soco.2007.25.5.687

Gawronski, B., & Bodenhausen, G. V. (2014). Implicit and explicit evaluation: A brief review of the Associative-Propositional Evaluation model. *Social and Personality Psychology Compass*, *8*(8), 448–462. https://doi.org/10.1111/spc3.12124

Gawronski, B., & De Houwer, J. (2014). Implicit measures in social and personality psychology. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2nd ed., pp. 283–310). New York, NY: Cambridge University Press.

Gawronski, B., Morrison, M., Phills, C. E., & Galdi, S. (2017). Temporal stability of implicit and explicit measures: A longitudinal analysis. *Personality and Social Psychology Bulletin*, *43*(3), 300–312. https://doi.org/10.1177/0146167216684131

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Journal of Personality and Social Psychology*, *102*(1), 4–21. https://doi.org/10.1037/0033-295x.102.1.4

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., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197–216. https://doi.org/10.1037/h0087889

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

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

Hehman, E., Calanchini, J., Flake, J. K., & Leitner, J. B. (2019). Establishing construct validity evidence for regional measures of explicit and implicit racial bias. *Journal of Experimental Psychology: General*, *148*(6), 1022–1040. https://doi.org/10.1037/xge0000623

Hehman, E., Flake, J. K., & Calanchini, J. (2018). Disproportionate use of lethal force in policing is associated with regional racial biases of residents. *Social Psychological and Personality Science*, *9*(4), 393–401. https://doi.org/10.1177/1948550617711229

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

Jones, I. F., Young, S. G., & Claypool, H. M. (2011). Approaching the familiar: On the ability of mere exposure to direct approach and avoidance behavior. *Motivation and Emotion*, *35*(4), 383–392. https://doi.org/10.1007/s11031-011-9228-7

Jost, J. T., Pelham, B. W., & Carvallo, M. R. (2002). Non-conscious forms of system justification: Implicit and behavioral preferences for higher status groups. *Journal of Experimental Social Psychology*, *38*(6), 586–602. https://doi.org/10.1016/S0022-1031(02)00505-X

Judd, C. M., & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review*, *100*(1), 109–128. https://doi.org/10.1037/0033-295X.100.1.109

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. https://doi.org/10.1037/a0028347

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

Krieglmeyer, R., & Deutsch, R. (2010). Comparing measures of approach–avoidance be- haviour: The manikin task vs. two versions of the joystick task. *Cognition & Emotion*, *24*(5), 810–828. http://doi.org/10.1080/02699930903047298

Krosnick, J. A., & Judd, C. M. (1982). Transitions in social influence at adolescence: Who induces cigarette smoking? *Developmental Psychology*, *18*(3), 359–368. https://doi.org/10.1037//0012-1649.18.3.359

Krosnick, J. A., Judd, C. M., & Wittenbrink, B. (2005). The measurement of attitudes. In D. Albarracín, B. T. Johnson & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 21–76). Mahwah, NJ: Erlbaum.

Kurdi, B., & Banaji, M. R. (2017). Reports of the death of the individual difference approach to implicit social cognition may be greatly exaggerated: A commentary on Payne, Vuletich, and Lundberg. *Psychological Inquiry*, *28*(4), 281–287. https://doi.org/10.1080/1047840x.2017.1373555

Kurdi, B., Seitchik, A. E., Axt, J. R., Carroll, T. J., Karapetyan, A., Kaushik, N., … Banaji, M. N. R. (2019). Relationship between the Implicit Association Test and intergroup behavior: A meta-analysis. *American Psychologist*, *74*(5), 569–586. https://doi.org/10.1037/amp0000364

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology*, *81*(5), 842–855. https://doi.org/10.1037/0022-3514.81.5.842

Lynott, D., Kansal, H., Connell, L., & O'Brien, K. (2012). Modelling the IAT: Implicit Association Test reflects shallow linguistic environment and

not deep personal attitudes. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *34*(34), 1948-1953.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Markman, A. B., & Brendl, C. M. (2005). Constraining theories of embodied cognition. *Psychological Science*, *16*(1), 6–10. https://doi.org/10.1111/j.0956-7976.2005.00772.x

Neumann, R., Hülsenbeck, K., & Seibt, B. (2004). Attitudes towards people with AIDS and avoidance behavior: Automatic and reflective bases of behavior. *Journal of Experimental Social Psychology*, *40*(4), 543–550. https://doi.org/10.1016/j.jesp.2003.10.006

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 516–524. https://doi.org/10.3758/BF03195598

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

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, *6*(1), 101–115. https://doi.org/10.1037/1089-2699.6.1.101

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

Oliveira, M., Garcia-Marques, T., Garcia-Marques, L., & Dotsch, R. (2019). Good to bad or bad to bad? What is the relationship between valence and the trait content of the big two? *European Journal of Social Psychology*, 1-21. https://doi.org/10.1002/ejsp.2618

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

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Orchard, J., & Price, J. (2017). County-level racial prejudice and the black-white gap in infant health outcomes. *Social Science & Medicine*, *181*, 191–198. https://doi.org/10.1016/j.socscimed.2017.03.036

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

Paladino, M.-P., & Castelli, L. (2008). On the immediate consequences of intergroup categorization: Activation of approach and avoidance motor behavior toward ingroup and outgroup members. *Personality and Social Psychology Bulletin*, *34*(6), 755–768. https://doi.org/10.1177/0146167208315155

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

Payne, B. K., Vuletich, H. A., & Brown-Iannuzzi, J. L. (2019). Historical roots of implicit bias in slavery. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(24), 11693–11698. https://doi.org/10.1073/pnas.1818816116

Payne, B. K., Vuletich, H. A., & Lundberg, K. B. (2017). The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychological Inquiry*, *28*(4), 233–248. https://doi.org/10.1080/1047840x.2017.1335568

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*(3), 811–832. https://doi.org/10.1037//0022-3514.75.3.811

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510–532. https://doi.org/10.1037/0033-2909.114.3.510

Richeson, J. A., Trawalter, S., & Shelton, J. N. (2005). African Americans' implicit racial attitudes and the depletion of executive function after interracial interactions. *Social Cognition*, *23*(4), 336–352. https://doi.org/10.1521/soco.2005.23.4.336

Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*(2), 105–120. https://doi.org/10.1016/j.jbtep.2006.10.001

Rougier, M., Muller, D., Ric, F., Alexopoulos, T., Batailler, C., Smeding, A., & Aubé, B. (2018). A new look at sensorimotor aspects in approach/avoidance tendencies: The role of visual whole-body movement information. *Journal of Experimental Social Psychology*, *76*, 42–53. https://doi.org/10.1016/j.jesp.2017.12.004

Rudman, L. A., Feinberg, J., & Fairchild, K. (2002). Minority members' implicit attitudes: Automatic ingroup bias as a function of group status. *Social Cognition*, *20*(4), 294–320. https://doi.org/10.1521/soco.20.4.294.19908

Seibt, B., Neumann, R., Nussinson, R., & Strack, F. (2008). Movement direction or change in distance? Self- and object-related approach–avoidance motions. *Journal of Experimental Social Psychology*, *44*(3), 713–720. https://doi.org/10.1016/j.jesp.2007.04.013

Shepherd, H. (2011). The cultural context of cognition: What the Implicit Association Test tells us about how culture works. *Sociological Forum*, *26*(1), 121–143. https://doi.org/10.1111/j.1573-7861.2010.01227.x

Stanley, D. A., Sokol-Hessner, P., Banaji, M. R., & Phelps, E. A. (2011). Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(19), 7710–7715. https://doi.org/10.1073/pnas.1014345108

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*(3), 220–247. https://doi.org/10.1207/s15327957pspr0803_1

Vaes, J., Paladino, M. P., Castelli, L., Leyens, J.-P., & Giovanazzi, A. (2003). On the behavioral consequences of infrahumanization: The implicit role of uniquely human emotions in intergroup relations. *Journal of Personality and Social Psychology*, *85*(6), 1016–1034. https://doi.org/10.1037/0022-3514.85.6.1016

Wentura, D., Rothermund, K., & Bak, P. (2000). Automatic vigilance: The attention-grabbing power of approach- and avoidance-related social information. *Journal of Personality and Social Psychology*, *78*(6), 1024–1037. https://doi.org/10.1037/0022-3514.78.6.1024

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, *143*(5), 2020–2045. https://doi.org/10.1037/xge0000014

Wiers, C. E., Kühn, S., Javadi, A. H., Korucuoglu, O., Wiers, R. W., … Bermpohl, F. (2013). Automatic approach bias towards smoking cues is present in smokers but not in ex-smokers. *Psychopharmacology (Berl)*, *229*(1), 187–197. https://doi.org/10.1007/s00213-013-3098-5

Wolsiefer, K., Westfall, J., & Judd, C. M. (2016). Modeling stimulus variation in three common implicit attitude tasks. *Behavior Research Methods*, *49*(4), 1193–1209. https://doi.org/10.3758/s13428-016-0779-0

Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology*, *10*(2), 109–120. https://doi.org/10.1016/0022-1031(74)90059-6

Zerhouni, O., Rougier, M., & Muller, D. (2016). "Who (really) is Charlie?" French cities with lower implicit prejudice toward Arabs demonstrated larger participation rates in Charlie Hebdo rallies. *International*

Review of Social Psychology, *29*(1), 69–76. https://doi.org/10.5334/irsp.50

Zhou, Y., Li, X., Zhang, M., Zhang, F., Zhu, C., & Shen, M. (2012). Behavioural approach tendencies to heroin-related stimuli in abstinent heroin abusers. *Psychopharmacology (Berl)*, *221*(1), 171–176. https://doi.org/10.1007/s00213-011-2557-0

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Rougier M, Muller D, Courset R, Smeding A, Devos T, Batailler C. Toward the use of approach/avoidance tendencies as attitude measures: Individual- and group-level variability of the ingroup bias. *Eur J Soc Psychol*. 2020;00:1–19. https://doi.org/10.1002/ejsp.2653