

# Advice-taking as an unobtrusive measure of prejudice

Thomas L. Webb

Published online: 17 June 2011  
© Psychonomic Society, Inc. 2011

**Abstract** A new unobtrusive measure of prejudice is proposed based on an advice-taking task. The computer-based task requires participants to find a token hidden behind one of two boxes. Prior to making their choice, however, someone (depicted by a name or a face) provides advice as to the token's location. An unobtrusive measure of prejudice is derived by manipulating the advice-giver's social group (e.g., male or female, Asian or White) and comparing the proportions of advice taken from each group. In [Experiment 1](#), although the participants were not aware of it, they took more advice from males than from females. In [Experiment 2](#), the relative proportion of advice taken from Asian versus White advice-givers correlated with responses to a news story pertaining to Asians. In [Experiment 3](#), the relative proportion of advice taken from Asian versus White advice-givers correlated with scores on the Implicit Association Test (IAT) and predicted discriminatory behaviour, as indexed by the lost e-mail technique, better than other measures. In [Experiment 4](#), scores on the advice task were uncontaminated by social desirability concerns and reactance and reflected the relative amounts of trust that people placed in different social groups. Taken together, these findings suggest that the advice task may be a useful tool for researchers seeking an unobtrusive measure of prejudice with predictive validity.

**Keywords** Prejudice · Attitudes · Stereotypes · Advice · Trust · Implicit association test

Prejudice cannot be measured simply by asking people about their attitudes towards particular social groups (e.g., Do you

like Asian people?). First, people may not recognise that they are prejudiced (Nisbett & Wilson, 1977). Second, self-reports may be unduly influenced by social desirability concerns (Rosenberg, 1969) and experimenter effects like reactance (Webb, Campbell, Schwartz, & Sechrest, 1966). Therefore, a substantial body of research has been dedicated to developing unobtrusive measures that index prejudice without participants being aware that this is what is being assessed (for a review, see Fazio & Olson, 2003).

There are, however, several compelling reasons for developing an additional measure. First, research often requires the repeated assessment of prejudice over multiple time points, especially when the impact of an intervention or change in policy is being evaluated. Given the problems associated with practice effects and shared method variance, a new measure of prejudice can provide an additional tool for researchers interested in measuring changes in prejudice over time. Second, correlations between measures of implicit attitude and behaviour have been less than convincing (e.g., Karpinski & Hilton, 2001), and a number of authors have argued that popular measures like the Implicit Association Test (IAT; Greenwald, McGhee, & Schwarz, 1998) might reflect associations that a person has been exposed to, in addition to—or instead of—the extent to which a person endorses or will use those evaluative associations (e.g., Olson & Fazio, 2004). Thus, an unobtrusive measure of how people actually respond and interact with people from different social groups would likely be a valuable addition to these measures. Finally, many measures are based on response latencies, and so require precision measurement and requisite hardware. In response to these concerns, the present research aimed to develop an ecologically valid, unobtrusive measure of prejudice that would reflect not only beliefs about particular social groups, but whether or not those beliefs are used by the individual to direct action.

---

T. L. Webb (✉)  
Department of Psychology, University of Sheffield,  
Western Bank,  
Sheffield S10 2TN, England, UK  
e-mail: t.webb@sheffield.ac.uk

## The advice-taking task

Mansell and Lam (2006) developed an advice-taking task to investigate decision-making processes in clinical disorders. The computer-based task required participants to find tokens that were hidden behind either a triangular or a square box. Prior to making their choice, however, a face appeared along with advice as to where the token was located. Participants were informed that they could choose whether to follow or to ignore the advice. Given that people draw trait inferences from the facial appearance of other people, even after very brief exposures (e.g., 100 ms; Willis & Todorov, 2006), the proposal of the present research is that manipulating characteristics of the advice-giver (e.g., male vs. female, young vs. old, Asian vs. White) could provide an unobtrusive measure of prejudice. Specifically, if Group A and Group B did not differ in the quality of their advice, yet a participant took more advice from Group A than from Group B, one might conclude that the participant was prejudiced against Group B.

## The present research

Four experiments investigated the efficacy of the advice task as an unobtrusive measure of prejudice. **Experiment 1** sought to demonstrate that characteristics of the advice-givers (namely, their gender) influenced the proportion of advice taken. **Experiment 1** also probed participants' awareness of the impact of the advice-giver's gender on their responses. **Experiment 2** investigated the concurrent validity of an Asian/White advice task (using names in place of faces) by comparing advice-taking scores with responses to a news story involving Asians. **Experiment 3** investigated the concurrent and predictive validity of an Asian/White advice task by comparing advice-taking scores with IAT scores, self-reported attitudes, and motivation to control prejudice (measured at the same time) and a measure of discriminatory behaviour (taken 1 week later). **Experiment 4** investigated the beliefs that underlie decisions on the advice task, with the aim of determining whether judgements reflect prejudice rather than undesirable influences like social desirability concerns or reactance.

## Experiment 1: Proof of concept

### Method

#### Participants

A group of 59 undergraduate students (39 females, mean age = 19.10 years) participated in return for course credit.

## Procedure

**Advice task** Participants sat at a computer and were presented the instructions,

In this task you will be shown two boxes on the screen—a red box and a blue box. A token is hidden in one of the boxes. Your task is to guess which box contains the token. Each time that you find a token you will be awarded one point. Your task is to get as many points as possible—you will be told your total at the end of the experiment.

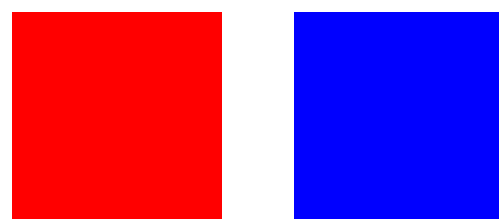
On the next screen, participants were told

However, there's a twist. Before you see the boxes a person will appear on the screen and give you advice about which box to choose. You can either make the decision by yourself or you can follow the person's advice. You should note that some people will give you better advice than others.

Twelve male and twelve female faces were selected from the Aberdeen image set (<http://pics.psych.stir.ac.uk/>) to represent advice-givers. A trial consisted of (1) a fixation cross for 1,000 ms, (2) a face and written advice ("Choose the red/blue box!") for 1,000 ms (see Fig. 1 for a screen shot), (3) a red and a blue box appear on the left and right of the screen, respectively, until the participant selects one, (4) feedback ("Correct! One point awarded" or "Incorrect") for 2,000 ms, and (5) an intertrial interval of 750 ms. There were 72 trials presented in a random order; each advice-



**Choose the red box!**



**Fig. 1** Screen shot from the advice task (Exp. 1)

**Table 1** Descriptive statistics and correlations between advice-taking scores and self-report measures (Exp. 1)

	Mean	SD	V1	V2	V3	V4	V5
V1. Prop. of advice taken from males	.67	.13					
V2. Prop. of advice taken from females	.59	.13	.48***				
V3. Advice-taking score (male – female)	.08	.13	.52***	–.50***			
V4. In this experiment, men gave better advice than women	4.90	1.72	–.17	–.16	–.01		
V5. Men tend to be more intelligent than women	3.85	2.07	–.10	–.28*	.17	.42**	
V6. I try not to judge people based on their appearance	5.10	1.64	–.02	–.15	.12	.15	–.09

\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$

giver was shown three times, twice giving correct advice and once giving incorrect advice.

*Self-report measures* At the end of the advice task, participants were asked to indicate the extent to which they agreed or disagreed (7-point scale) with eight statements. Embedded within these statements were three critical items: “Men tend to be more intelligent than women,” “In this experiment, men gave better advice than women,” and “I try not to judge people based on their appearance.”

## Results

Decisions that took  $> 2$  SDs longer than a participant’s mean response time (4.71% of total responses, mean cutoff = 859 ms) were removed from the analyses to ensure that decisions reflected relatively immediate responses to the advice. A 2 (gender of advice-giver [within subjects]: female vs. male) by 2 (gender of participant [between subjects]: female vs. male) repeated measures ANOVA was conducted, with proportion of advice taken as the dependent variable (see Table 1). The analyses revealed a significant main effect of the advice-giver’s gender,  $F(1, 57) = 14.67$ ,  $p < .001$ ,  $\eta^2 = .21$ . Participants were more likely to follow the advice of males ( $M = .67$ ,  $SD = .13$ ) than females ( $M = .59$ ,  $SD = .13$ ). The main effect of participant gender was nonsignificant,  $F(1, 57) = 0.77$ , n.s.,  $\eta^2 = .01$ , as was the interaction between the advice-giver’s gender and participant gender,  $F(1, 57) = 2.59$ , n.s.,  $\eta^2 = .04$ .<sup>1</sup>

Responses to the posttask questions suggested that participants did not perceive themselves as prejudiced, nor did they report favouring the advice of men over women

<sup>1</sup> Analyses of response latencies revealed no significant differences in the speeds of responses to male ( $M = 398$ ,  $SD = 175$ ) versus female ( $M = 390$ ,  $SD = 178$ ) advice-givers,  $F(1, 58) = 0.96$ , n.s.,  $\eta^2 = .02$ . However, participants were, on average, faster to accept ( $M = 383$ ,  $SD = 167$ ) than to reject ( $M = 414$ ,  $SD = 198$ ) advice,  $F(1, 58) = 7.95$ ,  $p < .01$ ,  $\eta^2 = .12$ .

(see Table 1). In order to investigate the relationship between the relative proportion of advice taken and self-reported gender prejudice, an advice-taking score was computed for each participant by subtracting the proportion of advice taken from females from the proportion of advice taken from males. This measure was then correlated with responses to the statements. Consistent with the idea that participants were not aware that they were displaying gender prejudice on the task, advice-taking scores did not correlate with any of the self-report items ( $r_s = .17$ ,  $.12$ , and  $-.01$ , respectively, all n.s.).

## Experiment 2: Concurrent validity of an advice task using names

**Experiment 2** investigated the concurrent validity of an Asian/White advice task by comparing advice-taking scores with responses to a news story involving Asians. The advice task was slightly modified to use names in place of faces—a useful modification if obtaining pictures of the relevant social group is difficult.

### Method

#### Participants

A group of 61 White British undergraduate students participated in return for course credit.<sup>2</sup>

#### Procedure

*Advice task* The advice task was identical to that used in **Experiment 1**, except that advice-givers were represented either by typically Asian names (Hussein, Youssef, Ahmed, Abdul), typically White names (John, Alex, Steven, Matthew, James, Michael, Thomas, Daniel, David, Harry, Chris, Andrew), or filler names that did

<sup>2</sup> Participants’ age and gender were not recorded in this experiment.

**Table 2** Descriptive statistics and correlations between advice-taking scores, the modified Modern Racism Scale, and responses to the news story (Exp. 2)

	Mean	SD	V1	V2	V3	V4	V5	V6	V7	V8
V1. Prop. of advice taken from Asians	.74	.19								
V2. Prop. of advice taken from Whites	.70	.15	.59***							
V3. Prop. of advice taken from filler names	.71	.14	.53***	.85***						
V4. Advice-taking score (White–Asian)	−.03	.16	−.61***	.28*	.10					
V5. Modified Modern Racism Scale	2.52	0.86	.04	.20	.18	.15				
V6. The crime is serious	6.71	0.76	.11	−.14	−.07	−.27*	−.29*			
V7. The crime is common	4.03	1.46	−.10	.14	.18	.26*	.20	−.13		
V8. The punishment is suitable	5.06	1.80	−.01	−.27*	−.16	−.26*	−.10	.26*	−.13	
V9. The perpetrator(s) will continue	3.92	1.35	.08	.36**	.26*	.26*	.10	.09	.23	−.34**

V6 to V9 reflect responses to the news story. Items are paraphrased for the table. \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$

not reflect a single social group (Fionn, Rafael, Diego, Akiko, Jean-Pierre, Kichirou, Wayne, Jason). Thus, instead of the face and written advice, only written advice appeared—for instance, “Hussein says choose the red box!” Filler names and a smaller proportion of Asian than White names were used so as not to arouse suspicion as to the true nature of the task. As before, the task comprised 72 trials, with each advice-giver appearing three times; twice giving correct advice and once giving incorrect advice.

**Measuring responses to a news story** Following the advice task, participants were asked to read a news story—“Asian gets life for race riot murder” (Wainwright, 2001)—taken from *The Guardian* newspaper. The article described a racially motivated attack by Asian perpetrators (see Appendix for the complete article). A real newspaper article (rather than say, a fictional story) was used to foster ecological validity (Goodman, Webb, & Stewart, 2009). Once participants had read the article, they were asked to indicate the extent to which they agreed or disagreed with four statements: “The crime committed is a serious problem,” “This type of crime is very common,” “The punishment given is suitable for the crime,” and “The perpetrator(s) will continue with criminal activity” (7-point scales anchored by *strongly disagree* and *strongly agree*).

**Modified modern racism scale** Following the advice task, participants completed a version of the modern racism scale (McConahay, 1983) adapted to focus on prejudice against Asians in the U.K., rather than against African-Americans in the U.S. Participants indicated the extent to which they agreed or disagreed with five statements: “Over the past few years, Asians have gotten more economically than they deserve,” “Over the past few years, the government and news media have shown more respect for Asians than they deserve,” “Discrimination

against Asians is no longer a problem in the United Kingdom,” “Asians are getting too demanding in their push for equal rights,” and “Asians should not push themselves where they are not wanted” (7-point scales anchored by *strongly disagree* and *strongly agree*). The five items were internally consistent (Cronbach’s alpha = .71) and were combined for analyses.

## Results

Advice-taking scores were computed for each participant by subtracting the proportion of advice taken from Asians from the proportion of advice taken from Whites. In support of the concurrent validity of the advice task, the relative proportion of advice taken from Whites versus Asians was significantly correlated with all four responses to the news story. A relative preference for advice from White over Asian advice-givers on the advice task was negatively correlated with the beliefs that the race-related attack described in the news story was serious ( $r = -.27$ ,  $p < .05$ ) and that the punishment was suitable ( $r = -.26$ ,  $p < .05$ ), and positively correlated with the beliefs that the crime was common ( $r = .26$ ,  $p < .05$ ) and that the perpetrator would likely continue with criminal activity ( $r = .26$ ,  $p < .05$ ) (see Table 2). In contrast, the modified Modern Racism Scale only correlated with one of the four responses to the news story, the belief that the race-related attack described in the news story was serious ( $r = -.29$ ,  $p < .05$ ). Advice-taking scores were only modestly (and nonsignificantly) correlated with scores on the Modern Racism Scale ( $r = .15$ , n.s.).<sup>3</sup>

<sup>3</sup> Analyses of response latencies revealed no significant differences in the speed of responses to Asian ( $M = 741$ ,  $SD = 454$ ) versus White ( $M = 722$ ,  $SD = 471$ ) advice-givers,  $F(1, 60) = 0.17$ , n.s.,  $\eta^2 = .00$ . However, participants were, on average, faster to accept ( $M = 681$ ,  $SD = 410$ ) than to reject ( $M = 811$ ,  $SD = 457$ ) advice,  $F(1, 60) = 23.81$ ,  $p < .001$ ,  $\eta^2 = .28$ .

## Discussion

The findings of [Experiment 2](#) provide further support for the value of the advice task as an unobtrusive measure of prejudice. Specifically, scores on the advice task predicted responses to a news story more consistently than did a version of the Modern Racism Scale. The findings also demonstrate that names can be used to represent advice-givers in the advice-taking task.

### Experiment 3: Concurrent and predictive validity of the advice task

[Experiment 3](#) sought to build on the findings of [Experiment 2](#) to investigate the concurrent and predictive validity of the advice task. Specifically, [Experiment 3](#) compared the advice task with perhaps the most established measure of implicit attitudes (the IAT) as a predictor of prejudiced behaviour.

## Method

### Participants

A group of 101 White British undergraduate students (88 females, mean age = 19.18 years) participated in return for course credit.

### Procedure

**Advice task** The advice task was identical to that used in [Experiment 1](#), except that the advice-givers were 18 people who volunteered to have their photograph taken and who identified themselves as Asian ( $N = 6$ ) or White ( $N = 12$ ). The pictures depicted head and shoulders only on a white background. There were an equal proportions of males and females in each category. The task comprised 54 trials, with each advice-giver appearing three times: twice giving correct advice and once giving incorrect advice.

**Implicit association test** The IAT was used to measure implicit evaluations of Asian people. On each block of trials, participants were given two category labels (e.g., “Asian” and “pleasant”) and were instructed to classify words into each category by pressing the left key (E) or the right key (I), respectively. The Asian names were *Abdul, Hussein, Omar, Youssef, Ahmed, Rafiq, Saiid, Maliq, Farid, and Ali*. The control category names were all Scottish: *Hamish, Douglas, Duncan, Malcolm, Iver, Murdoch, Ewan, Alistair, Scott, and Callum* (this comparison category had been effectively used previously to measure attitudes towards Asians by Webb, Sheeran, & Pepper, [in press](#)). The pleasant words were *lucky, honour, gift, peace, rainbow, laughter, love, joy, and pleasure*. The unpleasant

words were *poison, grief, disaster, hatred, evil, bomb, injury, disease, filth, and terror*. The order of compatible and incompatible blocks was counterbalanced across participants.

**Self-report measures** Following the advice task, participants were asked to select one of four statements: “I tended to take more advice from White/Black/Asian people” (three separate statements) or “Peoples’ race made no difference to whether I took their advice or not.” After completing the IAT, participants rated how warm or cold they felt towards Asian people (9-point scale anchored by *very cold* and *very warm*) and selected a statement that best described their attitude towards Asian people (“I strongly/moderately/slightly prefer Scottish people to Asian people,” “I like Scottish people and Asian people the same,” “I slightly/moderately/strongly prefer Asian people to Scottish people”). These two items were combined to give a self-report measure of explicit attitudes towards Asian people ( $r = .43, p < .001$ ). Participants were also asked to indicate the extent to which they agreed or disagreed (7-point scale) with two further statements: “I am personally motivated by my beliefs to be non-prejudiced towards Asians” and “Because of my personal values, I believe that using stereotypes about Asians is wrong.” These two items were combined to give a measure of motivation to control prejudice towards Asian people ( $r = .30, p < .01$ ).

**Measure of discriminatory behaviour** The “lost e-mail technique” (Stern & Faber, 1997) was used to measure participants’ discriminatory behaviour towards Asian people (see Bushman & Bonacci, 2004). One week after the experiment, participants received an incorrectly addressed e-mail from the undergraduate tutor at their University. The e-mail was addressed to Irfan Patel (a typically Asian name) and concerned a piece of coursework that had been submitted incorrectly, explaining that the candidate had 24 h to resubmit the coursework or else they would be penalised. The dependent variable was whether or not the participants replied to the e-mail.

## Results

Advice-taking scores were computed for each participant by subtracting the proportion of advice taken from Asians from the proportion of advice taken from Whites. Decisions that took  $> 2$  SDs longer than a participant’s mean decision time (4.03% of total responses, mean cutoff = 2,256 ms) were removed from the analyses. The raw response latencies for each participant on the IAT task were treated in accordance with the improved IAT scoring algorithm

**Table 3** Descriptive statistics and correlations between advice-taking scores, Implicit Association Test (IAT) scores, self-report measures, and discriminatory behaviour (Exp. 3)

	Mean	SD	V1	V2	V3	V4	V5	V6	V7
V1. Prop. of advice taken from Asians	.60	.15							
V2. Prop. of advice taken from Whites	.61	.19	.64***						
V3. Advice-taking score (White – Asian)	.00	.15	-.64***	.19					
V4. IAT score	.31	.36	-.16	-.01	.20*				
V5. Self-reported attitude towards Asians	0.00 <sup>a</sup>	0.85	-.02	.03	.05	-.24*			
V6. Motivation to control prejudice towards Asians	5.94	0.97	-.04	-.07	-.02	-.17	.46***		
V7. Reply to e-mail	0.58	0.50	.08	-.13	-.24*	.00	-.05	-.04	

<sup>a</sup> The two questions measuring explicit attitude towards Asians were measured on different scales (1–9 vs. 1–7) and so were standardised before being combined. The original means indicated that most participants (66%) reported liking Scottish and Asian people equally ( $M = 3.62$ ,  $SD = 0.73$ ) and felt moderately warm towards Asian people ( $M = 6.52$ ,  $SD = 1.57$ ). \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$

(Greenwald, Nosek, & Banaji, 2003). In support of the concurrent and predictive validity of the advice task, the relative proportion of advice taken from Whites versus Asians were correlated with IAT scores ( $r = .20$ ,  $p < .05$ ) and whether or not participants responded to the lost e-mail ( $r = -.24$ ,  $p < .05$ ) (see Table 3). Correlations were of small-to-medium size according to Cohen's (1992) criteria. Scores on the advice task did not correlate with self-reported attitudes towards Asians ( $r = .05$ , n.s.) or motivation to control prejudice towards Asians ( $r = -.02$ , n.s.).<sup>4</sup>

In a logistic regression to predict whether or not participants would respond to the e-mail, only the advice-taking measure was predictive (Wald = 5.92,  $p < .05$ ). IAT scores, self-reported attitude towards Asians, and motivation to control prejudice did not predict responses to the lost e-mail (Walds = 0.45, 1.09, and 0.62, respectively, n.s.). The vast majority of participants (90%) reported that race made no difference to whether they took their advice or not, therefore suggesting that participants were not aware that they were displaying racial prejudice on the advice task, despite it being the best predictor of their responses to the lost e-mail.

## Discussion

The findings of Experiment 3 provide further support for the concurrent and predictive validity of the advice task as an unobtrusive measure of prejudice; the relative proportion of advice taken from Asian versus White advice-givers correlated with scores on the IAT and better predicted discriminatory behaviour as indexed by the lost e-mail technique. Correlations were, however, only of

small to medium magnitude (Cohen, 1992), suggesting that prejudice as indexed by the advice-taking task is likely to be only one of a number of factors determining responses. For example, one might speculate that whether or not a person responds to a lost e-mail is also likely to be a function of how busy they are when the e-mail arrives in their inbox. In short, the findings of Experiment 3 support the idea that the advice task is an ecologically valid, unobtrusive measure of prejudice that reflects not only attitudes towards a target group, but also whether or not those attitudes are used by the individual to direct action.

## Experiment 4: Beliefs underlying decisions on the advice task

The purpose of Experiment 4 was to investigate the beliefs that underlie decisions on the advice task. Whether or not we take advice from someone is likely to be influenced not only by our attitudes towards specific advice-givers and the social groups that they represent (e.g., female, Asian), but also attitudes about their trustworthiness (McGinnies & Ward, 1980; Van't Wout & Sanfey, 2008), perceived competence (McGinnies & Ward, 1980; Todorov, Mandisodza, Goren, & Hall, 2005), friendliness (Oehler & Kohlert, 2009), and similarity to "us" (Burger, Messian, Patel, del Prado, & Anderson, 2004; Gino, Shang, & Croson, 2009). Prejudice, then, manifested as the unwarranted preference for taking advice from members of one social group over another, may reflect one—or a number of—these beliefs. To investigate this hypothesis, Experiment 4 measured participants' attitudes towards each of six advice-givers, along with their perceived trustworthiness, expertise, friendliness, and similarity. Measures were aggregated across the two social groups that the advice-givers represented (males and females), and analyses investigated relations between these measures and the relative proportion of advice taken from males versus

<sup>4</sup> Analyses of response latencies revealed no significant differences in the speeds of responses to Asian ( $M = 730$ ,  $SD = 312$ ) versus White ( $M = 734$ ,  $SD = 326$ ) advice givers,  $F(1, 100) = 0.11$ , n.s.,  $\eta^2 = .00$ . However, participants were, on average, faster to accept ( $M = 715$ ,  $SD = 322$ ) than to reject ( $M = 794$ ,  $SD = 345$ ) advice,  $F(1, 99) = 17.81$ ,  $p < .001$ ,  $\eta^2 = .15$ .

females. **Experiment 4** also sought to rule out the possibility that extraneous variables such as social desirability concerns or reactance might influence decisions.

## Method

### Participants

A group of 112 White British undergraduate students (94 females, mean age = 21.13 years) participated in return for course credit.

### Procedure

**Advice task** The advice task was similar to that used in **Experiments 1** and **3**, except that the advice-givers were six people who volunteered to have their photograph taken; three of the advice-givers were female, and three were male. The task comprised 72 trials, with each advice-giver appearing 12 times: 8 times giving correct advice, and 4 times giving incorrect advice. Advice-taking scores were computed for each participant by subtracting the proportion of advice taken from females from the proportion of advice taken from males. Decisions that took  $> 2$  *SDs* longer than a participant's mean decision time (4.33% of total responses, mean cutoff = 1,800 ms) were removed from the analyses.

**Self-report measures** Before completing the advice task, participants were asked to complete a questionnaire concerning their perceptions of each of the advice-givers. Each questionnaire contained a small photo of the advice-giver at the top of the page followed by a series of ratings as described below.

**Attitude** was measured using the stem “How do you feel about this person?” followed by three 7-point scales anchored with the adjectives *negative–positive*, *cold–warm*, and *dislike–like*. Attitude towards males was computed by averaging responses across the three scales for each of the three male advice-givers (i.e., nine measures in total; Cronbach's alpha = .87). Attitude towards females was computed in a similar way by averaging responses across the three scales for each of the three female advice-givers (Cronbach's alpha = .87).

**Trustworthiness** was measured using the stem “This person looks:” followed by three 7-point scales anchored with adjectives adapted from Ohanian's (1990) scale: *untrustworthy–trustworthy*, *dishonest–honest*, and *insincere–sincere*. The perceived trustworthiness of males (Cronbach's alpha = .81) versus females (Cronbach's alpha = .77) was computed by averaging scores across male versus female advice-givers.

**Perceived competence** was measured using the stem “This person looks:” followed by three 7-point scales

anchored with the adjectives *unintelligent–intelligent*, *stupid–smart*, and *dumb–clever*. The perceived competence of males (Cronbach's alpha = .86) versus females (Cronbach's alpha = .81) was computed by averaging scores across male versus female advice-givers.

**Friendliness** was measured using the stem “This person looks:” followed by three 7-point scales anchored with the adjectives *unhelpful–helpful*, *unfriendly–friendly*, and *unapproachable–approachable*. The perceived friendliness of males (Cronbach's alpha = .81) versus females (Cronbach's alpha = .74) was computed by averaging scores across the three male and three female advice-givers.

**Similarity** was measured using the stem “This person looks:” followed by three 7-point scales anchored with adjectives adapted from Pahl and Eiser (2006): *dissimilar to me–similar to me*, *different to me–the same as me*, *unlike me–like me*. The perceived similarity of males (Cronbach's alpha = .87) versus females (Cronbach's alpha = .83) was computed by averaging scores across male versus female advice-givers.

**Reactance** was measured with four items: “The advice seemed arbitrary to me, so I tried to do the opposite,” “Who was giving the advice influenced whether I took it or not,” “I did not really consider who gave the advice” (recoded), and “Some people seemed to give better advice than others.” The items formed a reliable scale (Cronbach's alpha = .70).

Finally, **social desirability concerns** were measured using the 33 items from the Marlowe–Crowne social desirability scale (Crowne & Marlowe, 1960). The scale contains 33 statements concerning personal attitudes or traits (e.g., “I have almost never felt the urge to tell someone off,” “I'm always willing to admit it when I make a mistake”). Participants simply indicated whether each was true or false of them.

## Results

Table 4 provides the descriptive statistics and correlations between all measures. In general, participants reported moderately positive attitudes towards both male and female advice-givers (*Ms* = 4.68 and 4.86, *SDs* = 0.78 and 0.73, respectively) and found both groups of advice-givers to be moderately trustworthy (*Ms* = 4.53 and 4.83, *SDs* = 0.68 and 0.59, respectively), competent (*Ms* = 4.73 and 4.76, *SDs* = 0.66 and 0.63, respectively), and friendly (*Ms* = 4.73 and 5.01, *SDs* = 0.76 and 0.62, respectively). Participants felt slightly less similar to the advice-givers (i.e., mean scores were below the midpoint), especially towards male advice-givers (*Ms* = 2.93 and 3.71, *SDs* = 1.00 and 0.84, for male and female advice-givers, respectively). Participants reported low levels of reactance (*M* = 2.73, *SD* = 1.32) and moderate social desirability concerns (*M* = 16.28, *SD* =

**Table 4** Descriptive statistics and correlations between advice-taking scores and measures of attitude, trustworthiness, competence, friendliness, similarity, reactance, and social desirability concerns (Exp. 4)

	Mean	SD	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20	
V1. Prop. of advice taken from males	.94	.09																					
V2. Prop. of advice taken from females	.94	.08	.81***																				
V3. Advice-taking score (males – females)	.00	.05	.38***	-.24*																			
V4. Attitude towards males	4.68	0.78	.09	.04	.08																		
V5. Attitude towards females	4.86	0.73	-.11	-.08	-.06	.70***																	
V6. Relative attitude (males – females)	-0.17	0.59	.25**	.15	.18	.47***	-.31**																
V7. Trustworthiness of males	4.53	0.68	.08	-.05	-.21	.60***	.34***	.37***															
V8. Trustworthiness of females	4.83	0.59	-.17	-.10	-.12	.31**	.52***	-.23*	.48***														
V9. Relative trustworthiness (males – females)	-0.30	0.65	.24*	.04	.33***	.34***	-.12	.60***	.60***	-.41***													
V10. Competence of males	4.73	0.66	-.01	-.08	.11	.54***	.44***	.17	.71***	.46***	.33***												
V11. Competence of females	4.76	0.63	-.05	-.03	-.04	.38***	.50***	-.11	.53***	.71***	-.09	.58***											
V12. Relative competence (males – females)	-0.03	0.59	.04	-.07	.17	.19*	-.05	.31**	.22*	-.25**	.46***	-.42***											
V13. Friendliness of males	4.73	0.76	.02	-.08	.15	.67***	.47***	.31**	.75***	.41***	.41***	.70***	.51***	.23*									
V14. Friendliness of females	5.01	0.62	.01	.04	-.06	.48***	.64***	-.16	.46***	.71***	-.17	.52***	.62***	-.09	.61								
V15. Relative friendliness (males – females)	-0.28	0.62	.01	-.14	.24*	.34***	-.07	.54***	.46***	-.21*	.68***	.34***	.00	.37***	.61***	-.25**							
V16. Similarity of males	2.93	1.00	.17	.13	.07	.28**	.09	.25**	.13	-.14	.26**	.08	-.12	.22*	.24*	-.08	.38***						
V17. Similarity of females	3.71	0.84	.02	-.04	.09	.14	.28**	-.15	.08	.11	-.02	.04	.25**	-.22*	.20*	.21*	.03	.14					
V18. Relative similarity (males – females)	-0.78	1.21	.13	.14	-.01	.13	-.12	.31**	.05	-.19*	.23*	.04	-.27**	.33***	.06	-.21*	.29**	.72***	-.58***				
V19. Reactance	2.73	1.32	-.43**	-.57***	.13	-.20	-.25	.07	.03	.02	.01	.18	.09	.10	-.15	-.34*	.15	-.21	-.27*	.01			
V20. Social desirability concerns	16.28	4.72	-.08	-.17	.12	.04	.08	-.05	-.11	.01	-.10	.01	-.01	.02	.01	.08	-.05	.10	-.04	-.11	.15		



4.72). Scores on the advice task (reflecting a relative preference for advice from male over female advice-givers) showed medium-sized (Cohen, 1992) positive correlations with relative trustworthiness ( $r = .33, p < .001$ ) and relative friendliness ( $r = .24, p < .05$ ). There were also small-sized (Cohen, 1992) positive correlations between scores on the advice task and explicit attitudes ( $r = .18, n.s.$ ) and perceived competence ( $r = .17, n.s.$ ), although neither of these correlations reached significance.

Linear multiple regression was conducted to investigate the beliefs that predict the relative proportion of advice taken from males versus females on the advice task. Advice-taking scores were regressed on attitude, trustworthiness, perceived competence, friendliness, similarity, reactance, and social desirability concerns. Table 5 shows the results of these analyses. The overall model was significant,  $F(7, 111) = 2.37, p < .05$ , and explained 14% of the variance in advice-taking scores. Inspection of the beta weights revealed that just one predictor—differential trust of males versus females—significantly predicted advice-taking scores ( $\beta = .32, p < .05$ ). The beta weight was positive, suggesting that participants who rated male advice-givers as more trustworthy than female advice-givers also tended to take more advice from males than from females. None of the other variables significantly predicted advice-taking scores ( $\beta s < .12, p s > .22$ ).<sup>5</sup>

## Discussion

Experiment 4 showed that scores on the advice task are uncontaminated by reactance or social desirability concerns and primarily reflect the relative amount of trust that participants place in people from different social groups—in this case, males versus females. The perceived friendliness, attitudes, and competence of male relative to female advice-givers all showed small positive correlations with advice-taking scores. However, of these beliefs, only the correlation between perceived friendliness and advice-taking scores reached significance, and none of these beliefs predicted advice-taking scores when considered alongside trust. The findings should not be interpreted as showing that attitudes towards the advice-givers are unimportant—indeed, in Experiment 3, advice-taking scores correlated with implicit attitudes towards the relevant social groups. However, in no experiment have advice-taking scores correlated with self-report measures of attitude towards particular social

**Table 5** Regression of advice-taking scores on attitude, trustworthiness, competence, friendliness, similarity, reactance, and social desirability concerns (Exp. 4)

Predictor	$\beta$	$t$	$p$
Relative attitude (males – females)	–.01	–0.10	.92
Relative trustworthiness (males – females)	.32	2.30	.05
Relative competence (males – females)	.04	0.38	.70
Relative friendliness (males – females)	.04	0.29	.77
Relative similarity (males – females)	–.11	–1.12	.27
Reactance	.07	0.75	.45
Social desirability concerns	.12	1.23	.22

groups. Taking the findings of Experiments 3 and 4 together, then, the advice task seems to provide an unobtrusive measure of prejudice that reflects peoples' implicit attitudes towards members of the relevant social groups, along with the extent to which members of such groups are trusted.

## General discussion

Four experiments investigated the potential of an advice task for providing an unobtrusive measure of prejudice. In Experiment 1, the gender of the advice-giver influenced the proportion of advice taken. This effect was consciously disavowed to the extent that participants' did not report taking more advice from males than from females. Experiments 2 and 3 focused on racial prejudice and provided evidence for the concurrent and predictive validity of the advice task. Specifically, the proportion of advice taken from Asian, relative to White, advice-givers correlated with responses to a news story describing a racially motivated attack (Exp. 2) and predicted discriminatory behaviour 1 week later (Exp. 3). In each case, the advice task predicted responses better than such established measures of prejudice as the Modern Racism Scale (McConahay, 1983) and the IAT (Greenwald et al., 1998). In Experiment 4, scores on the advice task were uncontaminated by reactance or social desirability concerns, and primarily reflected the relative amounts of trust that people place in different social groups.

One of the strengths of the advice task is that scores are likely to reflect implicit evaluations of the relevant social groups (e.g., an implicit preference for Whites over Asians), the relative amount of trust placed in members of these groups, as well as whether these evaluations and judgements are actually used by the individual to direct action (e.g., shall I take advice from this person?). In a sense then, the advice task falls between “pure” measures of implicit attitudes (e.g., the IAT) and “pure” measures of behaviour (e.g., social interaction as coded by McConnell & Leipold, 2001). This is a strength—for researchers interested in

<sup>5</sup> Analyses of response latencies revealed no significant differences in the speeds of responses to male ( $M = 511, SD = 179$ ) versus female ( $M = 514, SD = 191$ ) advice-givers,  $F(1, 109) = 0.15, n.s., \eta^2 = .00$ . However, participants were, on average, faster to accept ( $M = 505, SD = 173$ ) than to reject ( $M = 734, SD = 428$ ) advice,  $F(1, 90) = 45.15, p < .001, \eta^2 = .33$ .

measuring prejudice and predicting discriminatory behaviour, the advice task is relatively unique in providing an unobtrusive measure of responses to members of a social group that yet has ecological, predictive, and concurrent validity.

A number of issues may influence how the advice task is best employed in future research. First, like many other unobtrusive measures of prejudice, the advice task relies on comparative judgements; for instance, is the person more prejudiced against people from Group A than from Group B? Therefore, scores on the advice task could represent responses to Group A, responses to Group B, or both (see Brendl, Markman, & Messner, 2001, for a similar analysis of the IAT). This issue can, however, be addressed in the advice task by considering correlations between the proportion of advice taken from each group and the relative score. For example, in Experiments 2 and 3 the relative score was more strongly correlated with the proportion of advice taken from Asian, rather than from White, advice-givers, suggesting that scores primarily reflected participants' differing responses to Asians. In Experiments 1 and 4, the relative scores correlated with both the proportion of advice taken from males and the proportion of advice taken from females, suggesting that the scores reflected participants' differing responses to both social groups. Second, there are alternative ways of looking at the data from the advice task. For example, one might look simply at how participants respond the first time that they see an advice-giver (i.e., before responses are potentially contaminated by the quality of the person's initial advice). Alternatively, one could capitalise on potential temporal dynamics within the task in order to investigate whether participants are quicker to distrust the advice from members of one social group than from another by looking at the likelihood that participants will reject an advice-givers subsequent advice when their initial advice proved wrong.<sup>6</sup> Finally, if obtaining images of the relevant social group proved difficult, Experiment 2 suggests that the task could be modified to use names in place of pictures.

Finally, it is worth noting two potential limitations in the use of the advice task. First, although Experiment 4 suggested that scores were unrelated to reactance (e.g., intentionally trying to do the opposite of what was advised) and social desirability concerns, participants in the present experiments were never informed of what the advice task was intended to measure prior to completing the task. This mild deception is probably imperative to the success of the measure—if participants are aware of what the advice task is intended to measure, they are likely to find it relatively easy to intentionally subvert, or fake, responses (for discussions of faking, see Fiedler & Bluemke, 2005; Steffens, 2004).

Second, scores on the advice task are only likely to reflect prejudice when participants construe advice-givers in terms of the relevant social category (Ma & Correll, 2011; Olson & Fazio, 2003). Although categorisation is relatively automatic, it would be interesting to consider how multiple categorisation (Crisp & Hewstone, 2007) and manipulations that promote novel categorisations (e.g., Crisp, Walsh, & Hewstone, 2006) influence responses.

**Author Note** The author is grateful to Donna Collins, Alison Murphy, James Nicoll, Stephanie Parry, Elizabeth Tane, Nikita Woodcock, and Richard Woodward for assistance with data collection, and to Warren Mansell for providing additional information about his advice task.

## Appendix 1

### News Story used in Experiment 2

#### *Asian gets life for race riot murder*

A man who shot dead a student after cornering him during a street riot was jailed for life yesterday for what a judge called “premeditated and racially motivated murder”. Safdar Khan, 23, blasted Dester Coleman in the back as he tried to take refuge in a café from a machete and hammer-wielding mob.

In a five week trial at Bradford crown court, horrific scenes were described as Mr Khan and up to 100 other men, of predominantly Asian-origin, took revenge for insults during an earlier pub argument. The mob brandished “almost every sort of weapon” according to Mr Justice Gregson, as they chased a much smaller Afro-Caribbean group through the streets of Manningham.

Mr Coleman, who was 26 and on a computer studies course in Bradford, died from his injuries in the nearby Young Lions café which was surrounded and stoned by the mob. Detectives on West Yorkshire police's murder inquiry said he was an innocent victim of the violence and had not been involved in earlier confrontations which sparked it.

The court heard that bricks had smashed through windows before police and paramedics arrived, but attempts to resuscitate the dying man failed. The crowd shouted racist abuse during the attack and then ransacked the neighbouring Tote, making off with £1,400.

Khan, of Bradford was also given the maximum five year sentence for violent disorder, to run concurrently, and three other men involved in the violence were jailed. Mohammed Raja, 22, and Mohammed Shaffi, 26, both from Bradford; and Adelson Saws, 20, of Chapeltown, Leeds, were all sentenced to five years for violent disorder. Raja was also given a concurrent sentence of

<sup>6</sup> The author thanks Russell Fazio and an anonymous reviewer for this suggestion.

five years for robbery. Yousaf Khan, 26, also of Bradford, received three-and-a-half years for violent disorder.

## References

- Brendl, C. M., Markman, A. B., & Messner, C. (2001). How do indirect measures of evaluation work? Evaluating the inference of prejudice in the Implicit Association Test. *Journal of Personality and Social Psychology, 81*, 760–773. doi:10.1037/0022-3514.81.5.760
- Burger, J., Messian, N., Patel, S., del Prado, A., & Anderson, C. (2004). What a coincidence! The effects of incidental similarity on compliance. *Personality and Social Psychology Bulletin, 30*, 35–43. doi:10.1177/0146167203258838
- Bushman, B. J., & Bonacci, A. M. (2004). You've got mail: Using e-mail to examine the effect of prejudiced attitudes on discrimination against Arabs. *Journal of Experimental Social Psychology, 40*, 753–759. doi:10.1016/j.jesp.2004.02.001
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.
- Crisp, R. J., & Hewstone, M. (2007). Multiple social categorization. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 39, pp. 163–254). San Diego, CA: Elsevier Academic Press. doi:10.1016/S0065-2601(06)39004-1
- Crisp, R. J., Walsh, J., & Hewstone, M. (2006). Crossed categorization in common ingroup contexts. *Personality and Social Psychology Bulletin, 32*, 1204–1218. doi:10.1177/0146167206289409
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349–354.
- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology, 54*, 297–327. doi:10.1146/annurev.psych.54.101601.145225
- Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Tests. *Basic and Applied Social Psychology, 27*, 307–316. doi:10.1207/s15324834basps2704\_3
- Gino, F., Shang, J., & Croson, R. (2009). The impact of information from similar or different advisors on judgment. *Organizational Behavior and Human Decision Processes, 108*, 287–302. doi:10.1016/j.obhdp.2008.08.002
- Goodman, R. L., Webb, T. L., & Stewart, A. J. (2009). Communicating stereotype-relevant information: Is factual information subject to the same communication biases as fictional information? *Personality and Social Psychology Bulletin, 35*, 836–852. doi:10.1177/0146167209334780
- Greenwald, A. G., McGhee, D. E., & Schwarz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464–1480. doi:0022-3514/98/\$3.00
- 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*, 197–216. doi:10.1037/0022-3514.85.2.197
- Karpinski, A., & Hilton, J. L. (2001). Attitudes and the Implicit Association Test. *Journal of Personality and Social Psychology, 81*, 774–788. doi:10.1037/0022-3514.81.5.774
- Ma, D. S., & Correll, J. (2011). Target prototypicality moderates racial bias in the decision to shoot. *Journal of Experimental Social Psychology, 47*, 391–396. doi:10.1016/j.jesp.2010.11.002
- Mansell, W., & Lam, D. (2006). “I won't do what you tell me!” Elevated mood and the assessment of advice-taking in euthymic bipolar I disorder. *Behavior Research and Therapy, 44*, 1781–1801. doi:10.1016/j.brat.2006.01.002
- McConahay, J. B. (1983). Modern racism and modern discrimination: The effects of race, racial attitudes, and context on simulated hiring decisions. *Personality and Social Psychology Bulletin, 9*, 551–558. doi:10.1177/0146167283094004
- McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology, 37*, 435–442. doi:10.1006/jesp.2000.1470
- McGinnies, E., & Ward, C. D. (1980). Better liked than right: Trustworthiness and expertise as factors in credibility. *Personality and Social Psychology Bulletin, 6*, 467–472. doi:10.1177/014616728063023
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we know: Verbal reports on mental processes. *Psychological Review, 84*, 231–259. doi:00006832-197705000-00001
- Oehler, A., & Kohlert, D. (2009). Financial advice giving and taking—Where are the market's self-healing powers and a functioning legal framework when we need them? *Journal of Consumer Policy, 32*, 91–116. doi:10.1007/s10603-009-9099-4
- Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising, 19*, 39–52.
- Olson, M. A., & Fazio, R. H. (2003). Relations between implicit measures of prejudice: What are we measuring? *Psychological Science, 14*, 636–639. doi:10.1046/j.0956-7976.2003.psci.1477.x
- 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*, 653–667.
- Pahl, S., & Eiser, R. (2006). The focus effect and self-positivity in ratings of self-other similarity and difference. *British Journal of Social Psychology, 15*, 107–116. doi:10.1348/014466605X49582
- Rosenberg, M. J. (1969). The conditions and consequences of evaluation apprehension. In R. Rosenthal & R. L. Rosnow (Eds.), *Artifact in behavioral research* (pp. 279–349). New York: Academic Press.
- Steffens, M. C. (2004). Is the Implicit Association Test immune to faking? *Experimental Psychology, 51*, 165–179. doi:10.1027/1618-3169.51.3.165
- Stern, S. E., & Faber, J. E. (1997). The lost email method: Milgram's lost letter technique in the age of the internet. *Behavior Research Methods, Instruments, & Computers, 29*, 260–263. doi:10.3758/BF03204823
- Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science, 308*, 1623–1626. doi:10.1126/science.1110589
- Van't Wout, M., & Sanfey, A. (2008). Friend or foe: The effect of implicit trustworthiness judgements in social decision-making. *Cognition, 108*, 796–803. doi:10.1016/j.cognition.2008.07.002
- Wainwright, M. (2001, June 5). Asian gets life for race riot murder. *The Guardian*, p. xx.
- Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences*. Chicago: Rand McNally.
- Webb, T. L., Sheeran, P., & Pepper, J. (in press). Gaining control over responses to implicit attitude tests: Implementation intentions engender fast responses on attitude-incongruent trials. *British Journal of Social Psychology*. doi:10.1348/014466610X532192
- Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*, 592–598. doi:10.1111/j.1467-9280.2006.01750.x