The Psychological Record (2021) 71:423–433
https://doi.org/10.1007/s40732-020-00427-w

ORIGINAL ARTICLE

Using the Implicit Relational Assessment Procedure (IRAP) to Examine Implicit Beauty Bias in the Context of Employability

Rachel Murphy[1] · Carol Murphy[1] · Michelle Kelly[2] · Bryan Roche[1]

Published online: 6 September 2020
© Association for Behavior Analysis International 2020

Abstract

Attractiveness bias has been well-documented in social domains, however, an investigation into the attractiveness-bias effect on employability has not been conducted using implicit measures. In Study 1 ($N = 24$) the Implicit Relational Assessment Procedure (IRAP) and a number of explicit measures (e.g., self-report questionnaires) were used to investigate the attractiveness-bias effect on employability using stimuli of high and low attractiveness. Results from Study 1 indicated that there was a significant bias in the direction of attractive-employable on explicit and implicit measures. In Study 2, ($N = 52$) these measures were used to investigate the attractiveness-bias effect on employability using stimuli of high and medium attractiveness. Results from Study 2 indicated that there was a significant bias in the direction of attractive-employable on explicit measures and a significant bias in the direction of attractive-employable and medium-attractive-unemployable on implicit measures. There was no effect of participant gender on *D*-scores for either study. Findings are discussed in relation to previous research and implications for the use of implicit measurement to measure attractiveness bias in the domain of employability.

Keywords implicit attitudes · IRAP · attractiveness bias · employability · RFT

Dion, Berscheid, and Walster (1972) provided one of the most influential quotes on research in physical attractiveness, that in our perceptions of other people "What is beautiful is good" (p. 285). This classic study linking physical attractiveness and positivity provides compelling evidence for the existence of attractiveness bias. In particular, that more attractive individuals are not only rated with preferable and favorable personality traits (e.g., talent, kindness, honesty, and intelligence) but also more successful life outcomes (e.g., income and marital success) based solely on the extent of their physical attractiveness (Schneider, Gruman, & Coutts, 2012). Following this influential and widely cited study, a large body of research has been conducted on the effects of an attractiveness bias, in particular, in social contexts and attractiveness-bias has been well-documented as a robust effect (Griffin & Langlois, 2006; Hosoda, Coats, Stone-Romero, & Backus, 1999; Langlois et al., 2000).

There has been less extensive research on the effects of attractiveness in the area of employment, but findings for the most part have confirmed that physical attractiveness bias does influence employment and hiring decisions (Shahani-Denning, 2003), and that attractive employees were more likely to be promoted to managerial positions and chosen for management training (Cash & Kilcullen, 1985). Research suggested that a physical attractiveness bias influences our perceptions of job candidates, and subsequently our perceptions of employee performance related to advancement (Schneider et al., 2012). Physically attractive individuals are more often seen as confident, intelligent, trustworthy, and competent, all highly desirable traits in a prospective employee.

The issue is not always straightforward in the context of employment, however, and research has suggested that gender is a relevant factor, in particular that attractive males are preferred over attractive females in hiring situations (Cash &

✉ Rachel Murphy
Murphy.rachel.94@gmail.com

Carol Murphy
Carol.a.murphy@mu.ie

Michelle Kelly
Michelle.Kelly@ncirl.ie

Bryan Roche
Bryan.t.roche@mu.ie

[1] National University of Ireland, Maynooth, Ireland

[2] National College of Ireland, Dublin, Ireland

Kilcullen, 1985). There have been reports also of a somewhat anomalous "beauty is beastly" effect (Heilman & Saruwatari, 1979), in some employment situations, where the more attractive candidates are less favored than their unattractive counterparts. Female employment candidates appear to be more vulnerable to this negative impact of attractiveness, in particular in the context of nontraditional or masculine job applications (Hosoda et al., 1999). Overall, findings suggest that physical attractiveness was almost always a benefit for men but was not always a benefit for women.

A possible limitation regarding the extant research literature on attractiveness bias is that it has mostly involved self-report measures such as questionnaires and rating scales, which are subject to well-documented limitations related to, for example, self-presentation effects (Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997), especially when issues are said to be socially "sensitive" (such as racial or weight bias). For example, even with an anonymous questionnaire, someone who holds a negative view of Black people may not honestly complete a self-report questionnaire or may downplay this socially undesirable bias (Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010). The problem of introspection is also relevant to self-report questionnaire data where some participants may not be aware of their own bias towards different social groups (Murphy, MacCarthaigh, & Barnes-Holmes, 2014).

To counteract disadvantages of self-report measures, a variety of procedures for evaluating so-called "implicit attitudes" have been developed, with the most popular being the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998). The IAT and similar measures are usually latency-based and require rapid participant responding in order to discourage participants from overthinking their responses and to gain their immediate reaction to the stimuli; for example, quicker responding in affirming attractive-employable compared to unattractive-employable would be deemed a demonstration of an implicit favorable bias toward attractiveness. Although there exists an on-going debate about the exact nature of implicit attitudes and how they influence behavior there is evidence that scores obtained on implicit measures (such as the IAT) reflect spontaneous and immediate responses, which may indicate that they are more automatic judgements than the considered and controlled behaviors typically seen on self-report measures (Friese, Hofmann, & Wänke, 2008). The Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes et al., 2006) is an implicit measure based on a behavior analytic theory of language that compares, under time pressure, response latencies when participants affirm relations presented onscreen that are either consistent or inconsistent with an assumed stereotype; mean reaction times are measured, and if it is shown that the participant group more rapidly affirms, for example *thin-positive* compared to *fat-positive* relations, this is deemed a pro-thin bias or IRAP effect (De Houwer & Moors, 2010). A distinct advantage of the IRAP in comparison to other implicit measures is that the direction of the bias (e.g., pro-thin or anti-fat) can be specified due to the four relational trial-type methodology. This means four relations are presented in IRAP trial-blocks, for example, attractive-employable-true/unattractive-employable-false (consistent with stereotype), unattractive-employable-true/attractive-employable-false (inconsistent with stereotype). The IRAP provides data for all four relations presented, which can help determine if the bias present is pro-attractive, anti-unattractive, some combination of the two, or no bias may be shown. This type of nuanced responding has been absent from extant attractiveness bias research to date, as noted by Griffin and Langlois (2006). The IRAP has been shown to be effective in detecting participants' implicit bias across multiple domains: pro-smoking attitudes (Vahey, Boles, & Barnes-Holmes, 2010), ageist attitudes (Cullen, Barnes-Holmes, Barnes-Holmes, & Stewart, 2009), and positive/negative self-esteem (Vahey, Barnes-Holmes, Barnes-Holmes, & Stewart, 2009). Recent research has suggested that various stimuli used in the IRAP methodology may not be as irrelevant as previously assumed (e.g., Maloney & Barnes-Holmes, 2016; Maloney, Foody, & Murphy, 2020); these findings do not negate research findings of an IRAP effect, but suggest caution regarding selection of IRAP stimuli, including response options. The response options selected in the current research were "true" and "false." Although Maloney and Barnes-Holmes (2016) suggest that relational response options (e.g., "Similar"/"Different") may facilitate detection of an IRAP effect more readily than natural language response options (e.g., "True"/"False"), coherence has also been found to be relevant (Finn, Barnes-Holmes, Hussey, & Graddy, 2016) and the True/False response options were selected because it was felt that these might readily cohere with the relations presented (e.g., attractive-positive-True) due to preexperimental learning.

The current research aimed to assess participant bias towards physical attractiveness in the context of employability using implicit—the IRAP (and explicit measures)—rating scales and questionnaires, such as the Interpersonal Judgement Scale (IJS; Byrne, 1971) and the Measures of Interpersonal Attraction (MIA; McCroskey & McCain, 1974). The current study is the first attempt to use implicit measurement in the context of physical attractiveness and employability; previous research in the area has mostly employed explicit self-report questionnaires (e.g., Griffin & Langlois, 2006) or behavioral tasks such as Curriculum Vita (CV) selection (e.g., Heilman & Saruwatari, 1979). Previous research in attractiveness-bias using the IRAP has focused primarily on physical appearance and evaluations of successfulness, but not employability (e.g., Murphy et al., 2014; Murphy, Hussey, Barnes-Holmes, & Kelly, 2015). This study aims to build upon the previous studies by utilizing the IRAP as a

method for measuring and determining the direction of attractiveness-bias in the domain of employability. The predictions for the current study were that a pro-attractive IRAP bias would be shown and participants' averaged response latencies would be shorter when affirming relations consistent with a pro-attractive stereotype compared to relations inconsistent with the stereotype; that explicit measures would indicate a bias favoring attractive individuals as more employable; and participants would select the mock CVs of attractive individuals more frequently compared to CVs of unattractive individuals.

Study 1

Method

Participants

Twenty-four individuals completed the Study 1 (11 male and 13 female), with an age range of 18–28. Six other individuals participated but their data were excluded because they failed to meet the predetermined criteria of the IRAP. Participants were undergraduate psychology students from National University of Ireland, Maynooth. University students were recruited based on the fact that college students may be representative of young middle-class adults who may form part of interview panels (Murphy et al., 2015).

Implicit measures The IRAP was presented on a standard laptop computer. The IRAP controlled all stimulus presentation and recordings of participant output (response latency data was recorded automatically). Each trial consisted of either an attractive or unattractive male or female image (see Figure 1) accompanied by an "good employee" or "bad employee" descriptive word (Table 1). Pictures were retrieved from the faceresearch.org website run by the University of Glasgow.

Explicit measures Participants were provided with a questionnaire booklet consisting of an employability and attractiveness measures, namely the Interpersonal Judgement Scale (IJS; Byrne, 1971) and a subscale (task attraction) of the Measures of Interpersonal Attraction scale (MIA; McCroskey & McCain, 1974). The IJS consists of four "distractor" questions presented before the two questions that measure interpersonal task attraction (e.g., how much would you like to work with this person?). Reliability measures for this scale were .85. High scores indicate high interpersonal task attraction, and that the person would be good to work with. In relation to attractive/unattractive stimuli (as used in the IRAP procedure), participants were required to indicate their agreement/disagreement with each question using a 5-point scale. In addition, two Likert scales were presented for each picture using the words "attractive" and "employable," and participants scored each picture by circling a number from -3 (very unattractive/unemployable) to +3 (very attractive/employable).

The MIA task attraction subscale consists of six statements that relate to working with an individual (e.g., He/she is a typical goof-off when assigned a job to do; I have confidence in his/her ability to get the job done) and participants were asked to rate their agreement/ disagreement with each statement using a 5-point scale. Reliability measures for the task attraction subscale were .86. High scores indicate that the individual would be good to work with, whereas low scores indicate that they would not be good to work with.

Behavioral task A brief behavioral task was implemented to ascertain if participant behavior was predicted by the preceding measures. This task involved presenting forced choices of four mock CVs for a human resource manager job vacancy. The participant was presented with two CVs at a time that differed only insofar as one depicted an attractive applicant (a picture used with other measures) and the other depicted an unattractive applicant; participants were asked to make a choice between the CVs of who they would recommend to proceed to interview.

Procedure

All procedures were conducted in the experimental labs at the Department of Psychology, National University of Ireland, Maynooth. Each participant was given an information sheet and consent form for the study. Participants were instructed that they were free to end their participation at any stage during the experiment and that all their data would be examined at group level and would remain confidential.

Implicit measures Each IRAP task began with a set of verbal instructions given by the researcher that described the task by explaining the layout of the screen and the response options. Participants were told that on each trial, either a "good employee" word or "bad employee" word would appear in the center of the screen along with a picture (either an attractive image or an unattractive image) at the top of the screen. In addition, the response options "true" and "false" were presented in the bottom left- and right-hand side of the screen, respectively. Participants were informed that the response options would be selected by pressing either the "d" or "k" key on the keyboard. Participants were asked to respond as quickly and accurately as possible, with at least 75% correct responding with a response latency below 2,000 ms. An example of each trial type was provided to the participant, and participants were told that there were practice blocks provided until the criterion was met. They were informed that after each block, the instructions and correct answers would be reversed, requiring the opposite pattern of responding from the previous

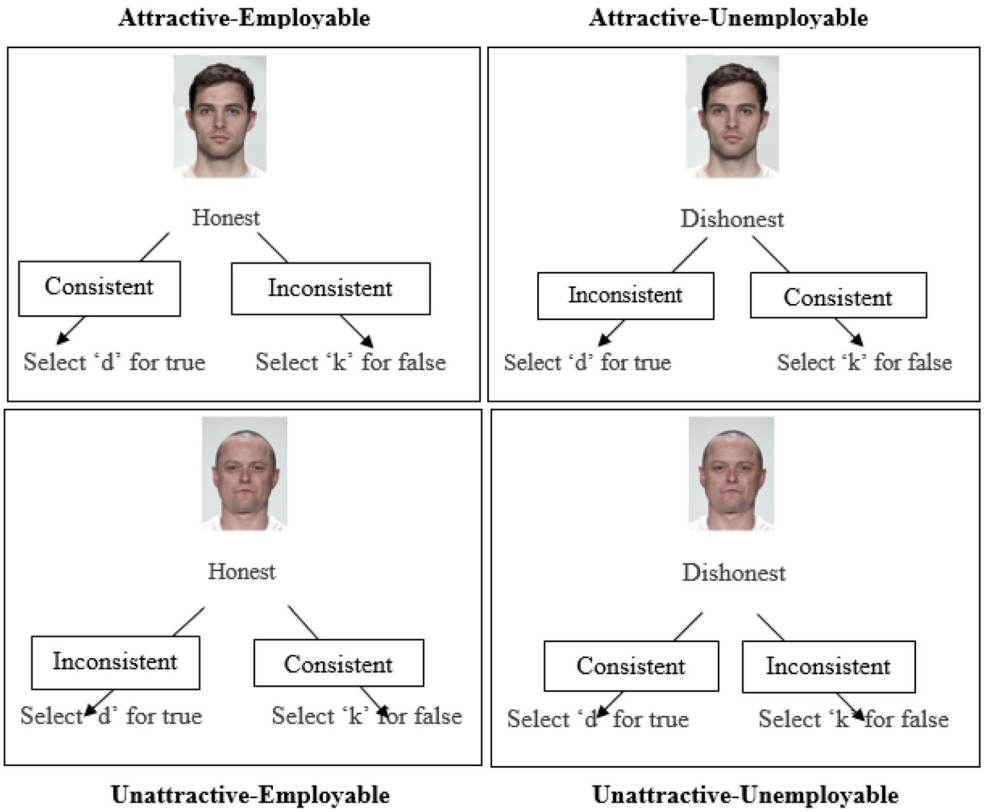

Fig. 1 Representations of four IRAP trial-types. The superimposed arrows and labels indicate what would be considered a bias towards attractive-employable (consistent) or a bias towards unattractive-employable (inconsistent) response for each trial type. (These are for illustration purposes and did not appear on screen)

block. Finally, participants were told that in some parts of the experiment they might be asked to respond in a way that did not reflect their beliefs but that this was part of the experiment.

Participants were required to choose one of these responses on each trial by pressing the "d" key for true and the "k" key for false. Response options remained static for the experiment. If the response option chosen was the option deemed correct for that trial, the screen cleared for a 400-ms interval, and the next trial was presented. If a participant chose the response option deemed incorrect (including pressing the wrong key), a red X appeared below the target stimulus and remained there until the correct option was chosen; once the correct option was chosen the computer then presented the 400-ms interval and continued as before. No feedback was presented about speed until the block was completed.

Before the participant was presented the first block of trials, a message appeared onscreen informing the participant that the first block was for practice purposes. The IRAP began with a block of trials that required responding that was consistent with pro-attractive-employable/pro-unattractive-unemployable. When the first trial-block was complete, the screen cleared, and a message appeared stating the rule to be followed for the next block of trials. This block was similar to the first (also for practice) except that the opposite responses were now required: pro-unattractive-employable and pro-attractive-unemployable. When both blocks were completed the

Table 1. Stimulus arrangements presented by the IRAP

Label 1	**Label 2**
Intelligent	Stupid
Trustworthy	Dishonest
Hard working	Lazy
Reliable	Unreliable
Approachable	Standoffish
Responsible	Irresponsible
Sample deemed consistent with Label 1	**Sample deemed consistent with Label 2**
Attractive faces	Unattractive faces
Response Option 1	**Response Option 2**
True	False

response accuracy and latency feedback for both blocks were presented, with instructions for continuing the experiment by pressing the spacebar on the keyboard.

If participants met the performance criteria for each of the practice trial-blocks, the IRAP program proceeded immediately to the test blocks (three pairs of blocks, three consistent and three inconsistent blocks). If the participant had not met the criteria, they were presented with a message reminding them of the criteria (above 80% accuracy and less than 2,000 ms) and a further pair of practice blocks. Participants had three pairs of practice blocks to reach the criteria. If they failed to do so they were thanked and debriefed. Participants who met the criteria progressed to the test blocks.

The three pairs of test blocks were similar to the practice blocks, except participants were informed that the block was a test and not practice. Once test blocks were completed, a message appeared informing participants that this part of the experiment was completed and to contact the researcher.

Explicit measures Participants were provided with a questionnaire booklet and instructed to complete the questionnaire by circling the number that best represented their response to each statement. Upon completion of the study, the participants were debriefed and thanked for their participation.

Results

Explicit Measures

The participant data ($N = 24$) collated from explicit measures were analyzed at the group level using paired sample *t*-tests and Wilcoxon signed rank tests. Bonferroni corrections were applied for all *t*-tests reported. All *t*-tests met the Bonferroni correction for *p*-values, $p < 0.012$. The overall results indicated that attractive images were rated as more attractive than unattractive images (statistically significant: $z = -4.29$, $p < .001$) and were rated more employable than unattractive images (also as statistically significant: $z = -3.79$, $p < .001$) across the three employability measures. For both the Measures of Interpersonal Attraction Scale and the Interpersonal Judgement Scale there was a statistically significant difference in scores for attractive and unattractive faces, $z = -3.48$, $p < .001$ and $z = -4.11$, $p < .001$, respectively. In addition to each of the explicit measures completed, a behavior task was presented to each participant (CV selection). Although it was predicted that CV selections would be in line with the explicit results (i.e., that CVs accompanied by attractive images would be chosen over CVs accompanied by unattractive images) data showed that there was no consistent participant selection of attractive CVs over unattractive CVs with 41% selecting equal numbers of attractive and unattractive CVs, 42% selected majority attractive CVs and 17% selected majority unattractive CVs.

Implicit Measure

The primary datum for the IRAP was response latency (from the first presentation of the stimuli in ms until the correct response was selected). Individual data was transformed into *D*-IRAP scores (see instructions in Hussey, Thompson, McEnteggart, Barnes-Holmes & Barnes-Holmes, 2015). A positive *D*-score was deemed pro-attractive employable whereas a negative *D*-score indicated pro-unattractive employable responding. However to compare directly across the trial types, and to produce a common axis and direction, the indicator signs for trial types 3 and 4 were inverted (e.g., reverse plus scores to minus, and minus scores to plus; see full explanatory details in Hussey et al., 2015). This inversion was to facilitate comparisons across trial types and does not alter the absolute value of IRAP *D*-scores. That is, scores above 0 indicate a bias towards attractive images as employable and scores below 0 indicate a bias towards unattractive as employable. This data transformation yielded an overall *D*-score that was positive, indicating an overall bias towards "attractive people are employable and unattractive people are unemployable." Conducting a single sample *t*-test revealed that this score (0.17) was statistically significantly different from 0, $t(23) = 3.88$, $p < .001$, eta square = 0.25.

The data from the 24 participants who completed the IRAP were included and the four mean *D*-IRAP scores for each trial are presented in Figure 2. Two of the trial types were statistically significantly different from 0: *Attractive-Employable* ($t(24) = 4.47$, $p < .001$, eta squared = .30); and *Attractive-Unemployable* ($t(24) = 3.92$, $p < .001$, eta squared = 0.25). The trial types *Unattractive-Employable* and *Unattractive-Unemployable* were not statistically significantly different from zero. These tests indicate that there were biases toward attractive images as employable and not unemployable, but no bias towards unattractive images in either direction. The 4-trial-type IRAP methodology allows for an examination of whether any bias shown is pro-attractive or anti-unattractive or some combination of both; the current trial-type data show that the bias present was pro-attractive in nature.

Gender Analysis

A 2x4 mixed between-within ANOVA was conducted to assess the impact of gender on participants' scores on the four IRAP trial types. The between-participant independent variable was gender, and the within-participant independent variable was trial type. The dependent variable was D-scores from the four IRAP trial types. There was no statistically significant interaction between gender and trial type, Wilks Lambda = .98, $F(3, 20) = .37$, $p = .78$, partial eta squared = .02, and no statistically significant main effect for gender ($p = 0.39$). This indicates that gender had no statistically significant

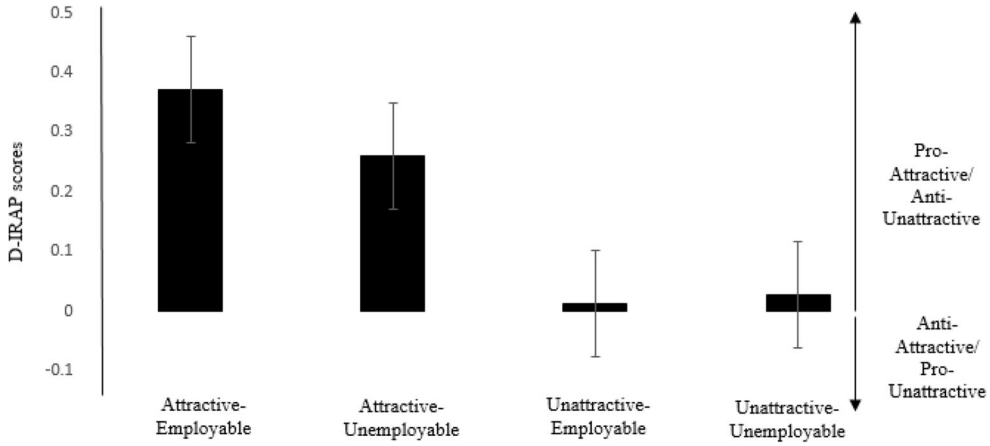

Fig. 2 Mean *D*-IRAP scores for each trial type.

effect on trial-type scores and there were no overall differences between males and females.

Implicit/Explicit Correlations

Each of the *D*-IRAP scores for the four trial types were entered into a correlation matrix with the four explicit measures. Of the resulting correlations, three were statistically significant. For the *Attractive-Unemployable* trial type the correlations indicated that an implicit bias towards images attractive as not-unemployable predicted that participants would rate attractive images as higher on Likert scales ($r = .41$, $p < 0.05$) and also rate unattractive images higher on the IJS ($r = .43$, $p < 0.05$). For the *Unattractive-Unemployable* trial type correlations indicated that as an implicit bias towards unattractive images as unemployable increased scores on the MIA decreased for unattractive images ($r = .47$, $p < 0.05$).

Prediction of CV Choice

To determine if the IRAP trial types had predictive effects on a behavioral task (CV choice between attractive/unattractive individuals), a multiple regression was conducted. This involved using the overall mean *D*-IRAP score (calculated by averaging the mean of the four-trial-type *D*-IRAP scores for each of the participants), where a negative score indicated an anti-attractive, pro-unattractive bias and a positive score indicated a pro-attractive, anti-unattractive bias. The Overall-*D* proved to be a nonsignificant predictor of CV choice, accounting for only 7% of the variance, $p > .05$. A further multiple regression was conducted, were each of the trial-type *D*-scores were entered into the model. Each of the trial-type *D*-scores also proved to be nonsignificant predictors of CV choice, with the model accounting for 8% of the variance, $p > .05$.

Summary

In summary, the explicit measures revealed that participants scored attractive images statistically significantly higher on the employability measures as well as on the Likert rating scale for employability. For the behavior task, there was an equal number of participants who chose a majority of attractive CVs and participants who chose an equal number of attractive and unattractive CVs. The overall *D*-IRAP score indicated that there was an attractive-employable and unattractive-unemployable bias. When individual trial type analysis was conducted a bias towards attractive images as employable and attractive images as not-unemployable across two trial types was indicated. Across the four trial types, the attractive-employable and the attractive-unemployable were statistically significant. Two multiple regressions were conducted to ascertain the influence that IRAP scores (both overall-*D* and *D*-scores per trial type) had on CV choice, however neither of the models were statistically significant in explaining the variance in responding for CV choice. Correlations between the implicit and explicit measures revealed three statistically significant correlations. For the attractive-unemployable trial type the correlations indicated that an implicit bias towards attractive images as not-unemployable predicted that participants would rate attractive images as higher (Likert scales) and also rate unattractive images higher (IJS). For the unattractive-unemployable trial-type correlations indicated that an implicit bias towards unattractive images as unemployable was related to increased scores on the MIA decreased. An ANOVA conducted to determine if gender of participants influenced scores on the four-IRAP trial-types *D*-scores indicated that there was no influence of gender on scores.

Study 2

The second study of the current research aimed to determine if the findings of Study 1 would be similar by using of images

individuals of medium attractiveness (rather than the low attractiveness used in Study 1) in the same population of undergraduate university students. The use of only either very high or very low attractiveness stimuli in physical attractiveness research was highlighted as a limitation in the available research, because little research has been conducted using more ambiguous images of stimuli of medium attractiveness (Langlois et al., 2000; Griffin & Langlois, 2006). In particular, using the same measures as used in Study 1, would participants rate attractive individuals as more employable than those of medium attractiveness and would the same directionality of the bias towards attractive images be found (i.e., pro-attractive rather than anti-medium attractive)?

Method

Participants

Fifty-two individuals completed the study (18 male and 34 female), with an age range of 18–34. Three other individuals participated in the study but were excluded because they failed to meet the predetermined criteria of the IRAP (75% accuracy and latency of less than 2,000 ms).

Apparatus and Stimuli

The research was conducted at the National University of Ireland, Maynooth, in the psychology departments testing labs. The same testing rooms and laptop used in Study 1 were used in Study 2. All materials used were the same as in Study 1, except that the unattractive stimuli were replaced with ambiguous/medium attractiveness stimuli.

Procedure

The procedure followed for the second study was identical to the procedure used in the first study. The only difference in Study 2 was that the unattractive IRAP images used in Study 1 were replaced with different stimuli in Study 2. Examples of the updated four-IRAP trial-types can be found in Figure 3. All other procedures and apparatus were identical to that of Study 1.

Results

Explicit Measures

The explicit measures were analyzed using paired sample *t*-tests and Wilcoxon signed rank tests. Bonferroni corrections were applied for all *t*-tests reported. The overall results of the explicit measures indicated attractive faces were rated as more attractive than the medium attractiveness faces and were rated as more employable across the three employability measures. For the attraction Likert scale, a paired sample *t*-test indicated that there was a statistically significant difference in scores for attractive and medium attractive faces, $t\,(51) = -17.37$, $p < .001$. Likewise, a Wilcoxon signed rank test revealed a statistically significant difference in scores for attractive and medium attractive faces for the employability Likert scale, $z = -5.83$, $p < .001$. For both the Measures of Interpersonal Attraction Scale and the IJS there was a statistically significant difference in scores for attractive and medium attractive faces, $z = -4.60$, $p < .001$ and $z = -5.75$, $p < .001$, respectively. For all scales, there was a statistically significantly higher score for attractive faces than medium attractive faces, indicating that attractive faces were scored as more attractive and more employable than medium attractive faces. All *t*-tests met the Bonferroni correction for *p*-values, $p < 0.012$. Data showed the participants had no consistent selection of attractive CVs over medium attractive CVs with 44% selecting equal numbers of medium attractive and attractive CVs, 46% selecting majority attractive CVs and 10% selecting majority medium attractive CVs.

Implicit Measures

IRAP data for the second study was analyzed using an identical procedure to Study 1. The data from the 52 participants who completed the IRAP were examined also at the level of trial type, and four mean *D*-IRAP scores for each trial type. A one-sample *t*-test showed that data for three of the four IRAP trial types were statistically significantly different from 0: *Attractive-Employable* ($t\,(51) = 10.14$, $p < 0.001$, eta = 0.5); *Attractive-Unemployable* ($t\,(51) = 5.60$, $p < 0.001$, eta = 0.24); and *Medium Attractive-Unemployable* ($t\,(51) = 4.08$, $p < 0.001$, eta = 0.14). The data for the remaining IRAP trial-type (*Medium Attractive-Employable)* was not statistically significantly different from zero. The results from the individual IRAP trial-types indicate the presence of a bias towards attractive images as employable and not-unemployable and a bias towards medium attractive images as unemployable (Figure 4).

Gender Analysis

A 2x4 mixed between-within ANOVA was conducted to assess the impact of gender on participants' scores on the four IRAP trial types. The between-participant independent variable was gender, and the within-participant independent variable was trial type. The dependent variable was *D*-scores. There was no statistically significant interaction between gender and trial type, Wilks Lambda = .86, $F\,(3, 48) = 1.11$, $p = .37$ partial, eta squared = .14, and no statistically significant main effect for gender ($p = 0.47$). This indicates that gender had no statistically significant effect on trial type scores and there were no overall differences between males and females. There was a statistically significant main effect for trial type (Wilks Lambda = .59, $F\,(3, 48) = 4.63$ $p < .05$). Post-hoc

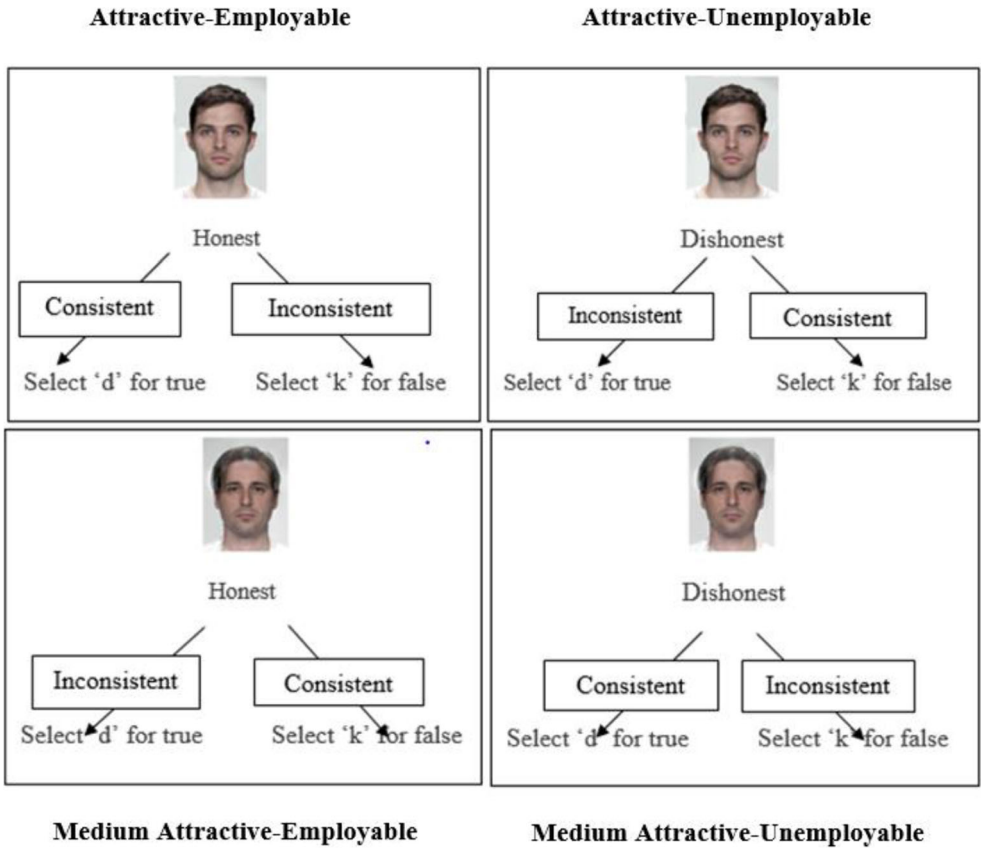

Fig. 3 Representations of four IRAP trial-types. The superimposed arrows and labels indicate what would be considered a bias towards attractive-employable (consistent) or a bias towards medium attractiveness-employable (inconsistent) response for each trial type. (These are for illustration purposes and did not appear on screen).

comparisons using paired-sample *t*-tests indicated that scores between all trial types were statistically significantly different from each other, at the Bonferroni alpha level ($p < .008$).and there were no overall differences between males and females.

Implicit/Explicit Correlations

Each of the *D*-IRAP scores for the four trial types were entered into a correlation matrix using Spearman's rho with the four explicit measures (8 scores, 4 for attractive faces and 4 for unattractive faces). Of the resulting correlations, none were statistically significant. In general, the scores on IRAP implicit measures did not predict scores on self-report measures.

Prediction of CV Choice

To determine if the IRAP trial types had predictive effects on a behavioral task (CV choice between attractive/medium attractive individuals), a multiple regression was conducted. This involved using the overall mean *D*-IRAP score (calculated by

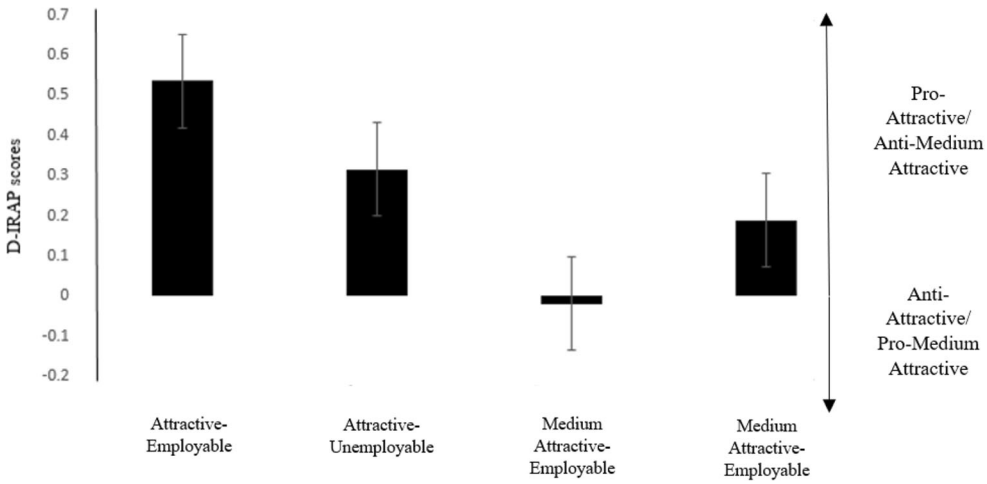

Fig. 4 Mean *D*-IRAP scores for the 4 IRAP trial types.

averaging the mean of the four-trial-type *D*-IRAP scores for each of the participants), where a negative score indicated an anti-attractive, pro-medium attractive bias and a positive score indicated a pro-attractive, anti-medium attractive bias. The Overall-*D* proved to be a nonsignificant predictor of CV choice, accounting for only 2% of the variance, $p > .05$. A further multiple regression was conducted, with each of the four-trial-types *D*-IRAP scores entered into the model. These also proved to be nonsignificant predictors of CV choice, with the model accounting for 2.7% of the variance, $p > .05$.

General Discussion

The current studies aimed to investigate the utility of using the IRAP as a measure of stereotype bias towards physical attractiveness in the context of employability. Although pro-attractive bias has been well-established in the self-report, explicit domain, little work has been done to date using implicit measures (Murphy et al., 2014). Overall, the results showed a statistically significant bias towards attractive individuals as more employable than unattractive individuals; this effect was found for comparisons between high- and low-rated attractiveness, and comparisons between high- and medium-rated attractiveness. The explicit measures revealed that participants scored attractive faces statistically significantly higher on the employability measures, both the IJS and the MIA, as well as on the Likert rating scale for employability. These results mirror results found in previous self-report studies conducted on attractiveness and employability. For the behavior task (a CV choice task based on attractive/unattractive individuals), only a small majority of participants choose three or more attractive CVs (out of a possible four). This result was unexpected, as previous studies indicated that attractive CVs were chosen consistently over unattractive CVs.

In Study 1 the overall *D*-IRAP score indicated that there was a statistically significant pro-attractive bias in participant responding, that is, responding per the consistent pro-attractive block was faster and more accurate than responding per the inconsistent pro-unattractive block. However, analysis of the four IRAP trial types indicate that the directionality of the bias was towards attractive-employable and attractive-not unemployable. In other words, participant responding was positive in regard to attractive images but neutral to unattractive images. In Study 2 however, there was an attractive-employable and no attractive-unemployable bias, as well as a medium-attractive-unemployable bias. As mentioned previously, the IRAP trial-type methodology facilitates an analysis of directionality of any bias shown, and in the current case the IRAP effect was pro-attractive but not anti-unattractive. An ANOVA revealed that there was no effect for gender on IRAP trial-type scores, which appears to support previous research that found sex of raters were not significant in whether unattractive individuals were rated as less employable than attractive individuals (Dipboye, Arvey, & Terpstra, 1977). Although findings in the current research suggest that participant gender was not influential, a potential limitation was that the research did not test whether participant gender might interact with the gender of the evaluee or target stimuli, as suggested in other research in the context of attractiveness bias and employability (Shahani-Denning, 2003). This limitation in the current research could be addressed in future IRAP research, because the methodology has been advanced such that the gender of the target stimuli can be recorded and included in analyses.

Correlational analyses revealed that there were three statistically significant correlations between the four IRAP trial types and the explicit measures in Study 1 and no statistically significant correlations in Study 2. In study 1, an attractive-not unemployable bias predicted higher ratings of attractive faces on the attractiveness scale and conversely, predicted higher ratings of unattractive faces on the IJS. An unattractive-unemployable bias predicted lower scores on the MIA task attraction subscale. Finally, regression analyses showed that the IRAP scores, both *D*-scores from the four trial types and the overall *D*-scores, were nonsignificant predictors of CV choice across the participants. However, this may be explained by the presentation effects of the CV choice task because they were presented after the explicit measures, so there may be an element of socially desirable responding in selecting both unattractive and attractive CVs, whereas in the IRAP there is no opportunity to engage in the lengthy response time found in self-report measures. However, if the IRAP does not predict responses in a behavioral task then it may limit its utility in measuring the attractiveness-bias effect. No correlations were found in the current study between explicit and implicit measures, and it remains unclear why this was the case, as at an intuitive level, the constructs measured appeared similar. This is not an uncommon result, nonetheless correlational tests of this type are of interest because it is possible that across a multitude of research findings we may begin to see patterns as to when correlations between findings using explicit and implicit methodologies might be expected. Low correlations between implicit and explicit data are thought to be related to motivation to self-report, method of measurement, independence of constructs (see meta-analysis of implicit-explicit correlations; Hofmann, Gawronski, Gschwendner, Huy, & Schmitt, 2005). Hofmann et al. found that correlations systematically increased as a function of increasing spontaneity of self-reports and increasing conceptual correspondence between measures.

A potential limitation of the current study is that the relatively small sample size may limit the generalizability of the results. However, it should be noted that Vahey, Nicholson, and Barnes-Holmes (2015) found sufficient statistical power with similar small samples. Likewise, the nature of the image

stimuli chosen is an important aspect of the current research. The current research employed images that were primarily Caucasian, which can limit the generalizability of the current results. These were chosen due to the make-up of the participant pool, which was composed of Caucasian individuals. However, it may be argued that the results obtained from the current study may not generalize across cultures and may be limited to cultures/countries where the majority of the population or sample is Caucasian. Further research should aim to determine the effects of using individuals of different races or participant samples comprising of different races. Second, the current research recruited undergraduate university students as the participant sample to measure the presence of possible biases towards attractive individuals. University students were recruited for the current research program based on the fact that college students may be representative of young middle-class adults who may form part of interview panels (Murphy et al., 2015). Nevertheless, it may be a possible limitation that the current research did not recruit individuals that regularly make such hiring decisions, such as HR individuals or managers. Future research should attempt to investigate the attractiveness bias using such a participant sample.

The results of the current study support the prediction that participants rated attractive individuals consistently more employable than unattractive individuals on both explicit and implicit measures. However, the IRAP methodology also allowed a more nuanced insight into the directionality of the bias. What is critical is that it revealed bias is in the direction of attractive individuals as more employable, rather than unattractive individuals as unemployable. This supports the IRAP to investigate influence of attractiveness and may, in the future, allow for further investigation of other relevant variables such as the influence of same-sex evaluation of attractiveness versus unattractiveness in the context of employability. Given that recent surveys have indicated that 60% of employers have searched for job candidates on their social network sites (Harris Poll on behalf of Career Builder, 2016), and sites such as LinkedIn, attractiveness bias may be affecting employment decisions even before selection of applicants for in-person interview.

The current study highlights the need for continued measurement of implicit attractiveness bias in the context of employment and other important domains, both for replication purposes and to assess potential effects of contextual variables and possible interventions. The current research extends both the existing research literature on implicit attractiveness bias, and research using the IRAP methodology, and is the first to do so in the current field. Furthermore, the current research adds a dimension that has not been investigated with the IRAP, in particular, the impact of different levels of attractiveness on implicit attractiveness bias and employability.

Availability of Data and Materials All data and materials from the current studies are available upon request from Rachel Murphy.

Compliance with Ethical Standards

Conflict of interest This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The research presented in the current study was part of a thesis submitted in part fulfilment of the requirements for the doctorate in psychological science, National University of Ireland, Maynooth.

Research involving human participants and/or animals The current research was approved by the social research ethics subcommittee in the Department of Psychology, National University of Ireland, Maynooth.

Informed consent Informed consent forms were presented to participants before the commencement of any part of the research. Forms included contact information and qualifications of both the researcher and supervisor. The forms outlined the expected duration of participation, the number of sessions required and how long the experiment should take. The informed consent forms also included information about how data will be stored and for how long. Accompanying the consent form was an information sheet, which explained the experiment in greater detail as well as data protection details: including how to remove data, request data, and where the experiment will take place. Procedures for withdrawing consent will also be addressed. Before each experiment session the researcher went through the consent form with the participant and informed them of their right to withdraw from the study at any time. Midway through the experiment session the researcher reminded them of their right to withdraw at any time.

References

Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, C., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the implicit relational assessment procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist, 32*, 169–177.

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2010). A sketch of the implicit relational assessment procedure (IRAP) and the relational elaboration and coherence (REC) model. *The Psychological Record, 60*, 527–542.

Byrne, D. (1971). *The attraction paradigm*. New York, NY: Academic Press.

Cash, T. F., & Kilcullen, R. N. (1985). The aye of the beholder: Susceptibility to sexism and beautyism in the evaluation of managerial candidates. *Journal of Applied Social Psychology, 18*, 591–605.

Cullen, C., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2009). The Implicit Relational Assessment Procedure (IRAP) and the malleability of ageist attitudes. *The Psychological Record, 59*, 591–620.

De Houwer, J., & Moors, A. (2010). Implicit measures: Similarities and differences. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications*. New York, NY: Guilford Press.

Dion, K. K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality & Social Psychology, 24*, 285–290.

Dipboye, R. L., Arvey, R. D., & Terpstra, D. E. (1977). Sex and physical attractiveness of raters and applicants as determinants of resume evaluations. *Journal of Applied Psychology, 62*, 288–294.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33*(5), 510–540.

Finn, M., Barnes-Holmes, D., Hussey, I., & Graddy, J. (2016). Exploring the behavioural dynamics of the implicit relational assessment procedure: The impact of three types of introductory rules. *The Psychological Record, 66*, 309–321.

Friese, M., Hofmann, W., & Wänke, M. (2008). When impulses take over: Moderated predictive validity of explicit and implicit attitude measures in predicting food choice and consumption behavior. *British Journal of Social Psychology, 46*, 397–419.

Greenwald, A. G., McGhee, A. D., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality & Social Psychology, 74*, 1464–1480.

Griffin, A. M., & Langlois, J. H. (2006). Stereotype directionality and attractiveness stereotyping: Is beauty good or is ugly bad? *Social Cognition, 24*(2), 187–206.

Harris Poll on behalf of Career Builder. (2016). Number of employers using social media to screen candidates has increased 500 percent over the last decade. Retrieved from http://www.careerbuilder.com/share/aboutus/pressreleasesdetail.aspx?ed=12%2F31%2F2016&id=pr945&sd=4%2F28%2F2016. Accessed 29 Mar 2017

Heilman, M. E., & Saruwatari, L. R. (1979). When beauty is beastly: The effects of appearance and sex on evaluations of job applicants for managerial and non-managerial jobs. *Organizational Behaviour & Human Performance, 23*, 360–372.

Hofmann, W., Gawronski, B., Gschwendner, T., Huy, L., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality & Social psychology Bulletin, 31*(10), 1369–1385.

Hosoda, M., Coats, G., Stone-Romero, E. F., & Backus, C. A. (1999, May). Who will fare better in employment-related decisions? A meta-analytic review of physical attractiveness research in work settings. Paper presented at the meeting of the Society of Industrial Organizational Psychology, Atlanta, Georgia.

Hussey, I., Thompson, M., McEnteggart, C., Barnes-Holmes, D., & Barnes-Holmes, Y. (2015). Interpreting and inverting with less cursing: A guide to interpreting IRAP data. *Journal of Contextual Behavioral Science, 4*, 157–162.

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larsen, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin, 126*, 390–423.

Maloney, E., & Barnes-Holmes, D. (2016). Exploring the behavioral dynamics of the Implicit Relational Assessment Procedure: The role of relational contextual cues versus relational coherence indicators as response options. *The Psychological Record, 66*(3), 1–9.

Maloney, E., Foody, M., & Murphy, C. (2020). Do response options in the Implicit Relational Assessment Procedure (IRAP) matter? A comparison of contextual relations versus relational coherence indicators. *The Psychological Record, 70*, 205–214.

McCroskey, J. C., & McCain, T. A. (1974). The measurement of interpersonal attraction. *Speech Monographs, 41*, 261–266.

Murphy, C., Hussey, T., Barnes-Holmes, D., & Kelly, M. E. (2015). The Implicit Relational Assessment Procedure (IRAP) and attractiveness bias. *Journal of Contextual Behavioral Science, 4*(4), 292–299.

Murphy, C., MacCarthaigh, S., & Barnes-Holmes, D. (2014). Implicit relational assessment procedure and attractiveness bias: Directionality of bias and influence of gender of participants. *International Journal of Psychology & Psychological Therapy, 14*(3), 333–351.

Schneider, F. W., Gruman, J. A., & Coutts, L. M. (2012). *Applied social psychology: Understanding and addressing social and practical problems*. London, UK: Sage.

Shahani-Denning, C. (2003). Physical attractiveness bias in hiring: What is beautiful is good. *Hofstra Horizons Spring Edition*, 14–18.

Vahey, N., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2009). A first test of Implicit Relational Assessment Procedure (IRAP) as a measure of self-esteem: Irish prisoner groups and university students. *The Psychological Record, 59*, 371–388.

Vahey, N., Boles, S., & Barnes-Holmes, D. (2010). Measuring adolescents smoking-related social identity preferences with the Implicit Relational Assessment Procedure (IRAP) for the first time: A starting point that explains later IRAP evolutions. *International Journal of Psychology & Psychological Therapy, 10*, 453–474.

Vahey, N., Nicholson, E., & Barnes-Holmes, D. (2015). A meta-analysis of criterion effects for the Implicit Relational Assessment Procedure (IRAP) in the clinical domain. *Journal of Behavior Therapy & Experimental Psychiatry, 48*, 59–65.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.