Using the Implicit Relational Assessment Procedure to examine implicit beauty-bias in the context of employability evaluations.

Rachel Murphy (B.A. Hons)

Supervisor: Dr Carol Murphy, Ph.D., BCBA-D

Co-Supervisors: Dr Michelle Kelly, Ph.D., BCBA-D & Dr Bryan Roche, Ph.D.

Thesis Presented in Part Fulfilment of the Requirements for the Doctorate in Psychological Science (Behaviour Analysis and Therapy), National University of Ireland, Maynooth.

Department of Psychology, Faculty of Science and Engineering

March 2017

Table of contents

Acknowledgements

I would like to thank my supervisors: Dr Carol Murphy, Dr Michelle Kelly and Dr Bryan Roche for all their help and invaluable guidance throughout the current research and the doctorate course. Their patience, support, advice and encouragement helped produce this thesis. I hope this thesis reflects all the teaching and effort they have given me.

My family and friends, for helping keep me sane: their encouragement and support, and for pretending to listen to me rant about my research. For Kevin, for keeping me sane and somewhat human over the course of this. For my parents, for knocking down any barriers, and for their constant support and belief in me. For everything you've done for me over the years, this thesis is dedicated to you.

All the participants who took part in this study, who gave me their time and effort. Without all of you this research would not have been possible – you were invaluable.

Lastly, thank you to all my fellow classmates for your support and help, and for sharing the load over the last three years.

Abstract

The current research aimed to examine the presence of attractiveness bias in ratings of employability by undergraduate university students using an implicit measure (the Implicit Relational Assessment Procedure; IRAP) and a number of explicit measures (Interpersonal Judgement Scale, Byrne, 1971; Measures of Interpersonal Attraction, McCroskey & McCain, 1974; and Likert scales). A number of explicit (self-report questionnaires etc.) have shown that attractiveness can be an influencing factor in important human social interactions. Limitations related to self-report measures, however, have been well documented, thus, the use of implicit measures such as the IRAP may extend research findings in this field.

In study 1, participants (*N*=24) were presented with an IRAP consisting of images of individuals of high and low attractiveness alongside the explicit measures and a behavioural task (a Curriculum Vitae choice task between unattractive and attractive individuals). Results revealed statistically significantly higher scores on employability measures for attractive images than for unattractive images, and a significant pro-beauty bias on implicit measures. There was no effect of gender on IRAP *D*-scores, and a multiple regression revealed no predictive power of the IRAP *D*-scores for the CV choice of participants. The procedure for Study 2 (*N*=52) was identical to study 1 with the exception of images of high and medium attractiveness used in the IRAP and explicit measures. Results from study 2 indicated similar results to study 1: there was statistically significantly higher scores on employability measures for high attractive images than for medium attractive images, and a statistically significant pro-beauty bias on implicit measures. However, the implicit measures also indicated that there was a statistically significant anti-unattractive bias present also. Again, there was no effect of gender on IRAP *D*-scores and no predictive power of the IRAP *D*-scores for participants' CV choice. Study 3 (*N*=28) aimed to investigate the 'beauty is beasty'

effect found for female applicants applying for executive jobs, using the same implicit measure as studies 1 and 2 and explicit measures taken from Heilman and Saruwatari (1979). Results from study 3 indicated statistically significantly higher ratings for male images in executive jobs and female images in nonexecutive jobs on explicit measures. On implicit measures, however, the IRAP indicated that there were statistically significant biases towards both male and female images in executive jobs. Again, there was no effect of gender on *D*-scores. 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.

CHAPTER 1

GENERAL INTRODUCTION

Using the Implicit Relational Assessment Procedure to examine implicit beauty-bias in the context of employability evaluations.

Dion, Berscheid and Walster (1972), in their seminal article, provided one of the most influential quotes on research in physical attractiveness, that is 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 the halo effect (Schneider, Gruman & Coutts, 2012), specifically that more attractive individuals are not only rated with more preferable and favourable personality traits (such as talent, kindness, honesty and intelligence) but also more successful life outcomes (such as income and martial success) based solely on the extent of their physical attractiveness. Following this influential and widely cited study, a large body of research has been conducted on the effects of an attractiveness bias in a number of contexts including legal contexts (Sigall & Ostrove, 1975); persuasion (Chaiken, 1979); likelihood of obtaining help (Benson, Karabenic & Lerner, 1976); as well as intelligence (Dion & Berscheid, 1971).

Thus it is evident that physical appearance and attractiveness are inarguably persuasive and powerful elements in the social context of nearly all individuals, and appear to have a high impact in many separate domains in the social world. School children mentioned people's appearance 590 times in a 22-hour observation period (Newman, Gabriel & Jones, 2001), attractive candidates received more votes compare to unattractive candidates in elections (Efran & Patterson, 1976) and teachers rated misbehaving children more favourably if they were physically attractive (Dion, 1972). Both adults and children prefer attractive to unattractive individuals, and there is evidence that even infants prefer the same attractive faces as adults (Langlois, Ritter, Roggman & Vaughn, 1991) and there is some evidence for consistent ratings of attractiveness across cultures (Langlois et al, 2000), where even cross-ethnic agreement of attractiveness were demonstrated. Physical attractiveness is an important

factor for choosing romantic partners, as within the normal ranges of physical attractiveness, highly attractive individuals are not only more preferred for romantic partners, but they are also more preferred as friends and seen as more likeable (Boyce, 1979). Attractive children are often judged as more adjusted, with higher interpersonal competence, as well as academic competence with teachers' ratings on how intelligent a child is, how far they would likely progress in school and even how popular they would be with peers significantly associated with the attractiveness of the child's photograph attached (Clifford & Walster, 1973). Even young children (as young as pre-schoolers) have stereotypes associated with appearance with both genders preferring pictures of attractive peers as potential friends, and even sometimes outright rejecting unattractive peers (Dion, 1973). Ramsey et al (2004) argued that infant's categorisation of facial attractiveness and their visual preferences for attractive faces serve as precursors to a development of such biases, with young girls (3-10 years old) referencing the attractiveness of their peers more often than boys of the same age (Cristofaro & Tamis-LeMonda, 2008). When asked to assign traits (both positive and negative) to peers, children consistently assigned attributes based on attractiveness, even when permitted to engage in non-forced choice tasks (where they could assign the attributes to one or the other, or both). Clearly, both appearance and attractiveness play an integral part of our social interactions, even from a very young age (Rennels & Langlois, 2014).

Attractiveness is linked to judgment of academic and occupational competence, not just social domains such as social appeal or interpersonal competence (Langlois et al, 2000). Attractiveness is also a strong predictor of assistance, with one study indicating that lost graduate school application forms with attached photographs of attractive individuals were more likely to be returned and sent than forms with attached photographs of unattractive individuals, even when an addressed, stamped envelope was provided as well (Benson, Karabenick & Lerner, 1976). Physical attractiveness also has influence in contexts in which it

would not be expected to play a large role, for example and perhaps surprisingly, there is evidence that physical attractiveness influences legal proceedings, not only in mock trials but in reviews of real world court proceedings (Stewart, 1980) where male defendants rated as attractive prior to their trials began were twice as likely to avoid jail sentences as unattractive defendants. There exists a substantial body of research that argues that physically attractive defendants were treated more leniently and received lesser sentences than unattractive defendants by mock jurors (Darby & Jeffers, 1988), especially when their attractiveness what not thought to be related to the crime (burglary). Interestingly, if attractiveness was seen to be related to the crime (swindle) the attractive defendants would receive a harsher treatment (Sigall & Ostrove, 1975). In civil cases, such as a staged negligence trial, plaintiffs who were more attractive than the defendants received almost twice the compensation of plaintiffs who were the more unattractive of the two (Kulka & Kessler, 1978). Even in relation to quite serious crimes physical attractiveness can influence judgements for jurors in mock trials. Jacobson and Popovich (1983) found that in mock rape trials physical attractiveness is an advantage to defendants, with more lenient sentences and decreased perceptions of guilt. Even the physical attractiveness of the mock victim was influential – with more attractive victims garnering more guilty verdicts and more severe sentences. The is an extension of the halo effect of attractiveness, with more physically attractive defendants and plaintiffs being seen as calmer, kinder, warmer, more sincere, and more intelligent (Castellow, Wuensch & Moore, 1990).

Eagly et al (1991) in their meta-analysis proposed that through the history of an individual, particularly in relation to the social context, certain personality traits, ideas about social competence and attributes become associated with physical attractiveness in order to create the halo effect so well documented in physical attractiveness research. This effect is argued to be the result of direct observations of attractive and unattractive individuals in the

social environment, as well as cultural representations of attractive and unattractive individuals in stories, films etc. Eagly et al argued that individuals observe that more attractive people receive more favourable reactions from others (such as popularity and preferential treatment) and observers may attribute this favourable reaction might be due not by their relative attractiveness but other positive personality traits. Feingold (1990) suggested that physical attractiveness does co-vary to some extent with social skills, social adjustment and an absence of shyness or anxiety, however, there is limited evidence for covariance between attractiveness and other personal attributes (such as intelligence, competency, kindness and sincerity). With regards to cultural influences, there is much less information about the extent of this influence on the development of biases. Eagly et al suggested that within stories attractiveness plays a part in identifying characters (with good characters being primarily attractive and vice versa) as well as advertising influence by pairing attractive individuals with desirable traits (Downs & Harrison, 1985). Although this makes intuitive sense, more information is required about the influence of the cultural context in order to make any statements about this effect on the development of physical attractiveness biases.

Some contextual factors may engender a negative side of the physical attractiveness bias, as there is some evidence to suggest that physical attractiveness may in fact cause people to infer negative characteristics about people. People often infer that attractive individuals are often more likely to have affairs (Brigham, 1980), or that their popularity makes it unnecessary for them to develop interpersonal qualities and sensitivity to main relationships in the same way as those who are not as attractive. Further, there is an argument for the ‘beauty is beastly’ effect, where physical attractiveness may be an impediment for women when applying for jobs typically seen as masculine (such as builders, truck drivers, etc.; see Heilman & Stopeck, 1985). First coined by Heilman and Saruwatari (1979) the effect was found that being attractive was a disadvantage only for female applicants when

applying for managerial roles, whereas being attractive for male applicants was consistently seen as an advantage for both managerial and non-managerial roles. Recent studies have found similar effects (Johnson et al, 2010; Paustian-Underdahl & Walker, 2016). Paustian-Underdahl and Walker (2016) found that the effect was strongest when female applicants were applying for jobs where attractiveness was not seen as a benefit, such as an IT manager and not an executive (where typically attractiveness is a benefit). In a performance appraisal context Shahani and Plumitallo (1993) found that supervisors were more likely to perceive attractive employees as failing due to a lack of motivation or effort, whereas unattractive employees were seen to fail due to bad luck or outside influences. Similarly, judgements made by the same sex may be more negative towards physically attractive individuals (Agthe, Spörrle & Maner, 2011). As such, it is difficult to explain such biases using the traditional model of observations in the real world and the cultural context of the individual as explained above.

A large amount of research has been conducted into the effects of physical attractiveness on employment decisions and hiring. Particularly given the high-stakes nature of employment decisions, it has been an important element of physical attractiveness research with the primary question being; Does physical attractiveness influence hiring decisions? The majority of research answers this overwhelmingly with yes; physical attractiveness bias does influence employment and hiring decisions (Shahani-Denning, 2003). Naturally if we are vulnerable to the stereotype that 'what is beautiful is good' regarding job applicants, a number of job positions with high public exposure will be filled by those that are physically attractive (McElroy & DeCarol, 1999), especially if it is believed that attractiveness will affect the profit of a company. Beehr and Gilmore (1982) found that in sales jobs and managerial jobs (both jobs which have an emphasis placed on face-to-face communication) physical attractiveness was an asset to hypothetical male interviewees. However, there still

remains some doubt to this finding, as another study using both hypothetical male and female interviewees found no such relationship between physical attractiveness and job type (Gilmore et al, 1986).

Even in job positions that would not require a high contact with the public, physical attractiveness still influences the hiring process (Watkins & Johnston, 2000), with considerable research supporting that the more attractive a candidate the more likely that individual will be hired. Further, even when provided with information regarding work experience and performance reviews the attractiveness bias was still very strongly in effect. Additionally, starting salaries are higher for more attractive candidates than less attractive peers (Jackson, 1983) and when progressing past the interview and selection process, physically attractive employees are still at an advantage over their unattractive peers, with attractive employees more likely to be promoted to managerial positions and chosen for management training (Cash & Kilcullen, 1985). A broader definition of physical attractiveness still yields a higher influence: with women with lower body-mass indexes and taller men receiving higher earnings and higher positions with organisations (Jæger, 2011). Within the US and Canada, attractive individuals are paid higher, with an average increase in pay of 12-14 percent over unattractive co-workers (Hammermesh & Parkel, 1994) and these advantages are consistent throughout the career life of attractive individuals. Frieze, Olson and Russel (1991) found in a longitudinal study of MBA graduates that the gap in income only increases over time, with every additional unit of attractiveness on a five-point scale related to an increase in pay over the lifetime. Surprisingly, even cultures outside of the western world, which typically have been found to be more collectivist and as such are hypothesised to be less influenced by physical attractiveness compared to individualist cultures, have been found to be susceptible to the physical attractiveness bias (Dion, Pak & Dion, 1990). The halo effect (Schneider, Gruman & Coutts, 2012) of physical attractiveness

highly influences our perceptions of job candidates, and even when hired our perceptions of employee performance and advancement. Physically attractive individuals are more often seen as confident, intelligent, trustworthy and competent, all highly desirable traits in a prospective employee. Although there is no evidence that suggests that these traits and physical attractiveness are linked in any strong and replicable way, there still is a persuasive view that attractive people do hold these traits more so than their unattractive peers and so are at an advantage in hiring decisions when decisions have to be made quickly with a limited amount of information about actual competence etc.

There appears to be a link between physical appearance of an individual and their gender, which was explained by Cash and Kilcullen (1985) in that physical attractiveness is preferred over unattractiveness, but that even within this, attractive males are preferred over attractive females in hiring. Similarly, intelligence is viewed to be higher in attractive males than unattractive females (Jackson, Hunter & Hodge, 1995). Female candidates also seem to be affected more by the "beauty is beastly" hypothesis (Heilman & Saruwatari, 1979), where physically attractive women are seen as inappropriate for traditionally masculine jobs – more so than unattractive females for the same job and attractive males for traditional feminine jobs. Similarly, when female managers received good performance evaluations this was more often put down to good luck or bias (Spencer & Taylor, 1988), which perhaps indicates that people are aware of the possible benefits that physical attractiveness can have. However, the relationship between sex and physical attractiveness may not always lead to a disadvantage for women. Shahani, Dipboye and Gehrlein (1983) found that for college interviews, interview scores for attractive women were consistently high, whereas for unattractive females and both attractive and unattractive males' physical appearance was only a benefit when their high school rank was also high. Hosodo. Coals, Stone-Romero and Backus (1999)

summed up the relationship between sex and physical attractiveness as nearly always a benefit for men and usually, but not always, a benefit for women.

Morrow (1990) put this negative effect of sex for women down to two observations: the assumption that women will be evaluated more based on looks than men in general day-to-day social interactions, and that especially in hiring studies done in the 1970s and 1980s there was a general preference for male candidates over women. How true each of these assumptions are remains a question for researchers of physical attractiveness, and does not explain what effect, if any, gender of evaluators have on the selection of candidates. Dipboye, Arvey and Terpstra (1977) investigated the effects of sex of raters on the hiring and evaluation of resumes in the domain of physical appearance. Similarly, they found that regardless of sex, attractive males contradictory fared best, with attractive females following. Female raters were just as likely to be biased against unattractive candidates as male raters, and those raters who were classified as physical unattractive were as susceptible to the same biases as those thought to be attractive. It has been noted however, that despite abundant research in the area of attractiveness-bias showing consistent robust effects, more nuanced research regarding gender effects is limited.

Another possible limitation regarding the extant research literature in attractiveness bias is the well documented major disadvantage to self-report measures such as questionnaires and rating scales, related to introspection and self-presentation effects (Dovidio, Kawakami, Johnson, Johnson & Howard, 1997). Especially in cases where responses are heavily influenced by the social context and when issues are said to be socially 'sensitive' (such as racial or weight biases) self-report measures may be complemented by additional measures of implicit bias. For example, it may be unlikely that someone who holds a very negative view of black people will honestly complete a self-report questionnaire without attempting to at least downplay this socially undesirable bias (Barnes-Holmes et al.,

2010). Further, the nature of the self-report measures may obscure biases that the participant cannot discriminate through introspection (Murphy, MacCarthaigh, Barnes-Holes, 2014).

In the specific domain of biases towards appearance and physical attractiveness Feingold (1990) speculated that results on attractiveness biases may be affected by the greater influence of self-presentational concerns among women over men. Specifically, ratings of attractive individuals that appear to be preferential over ratings of unattractive individuals may be perceived as unfair or shallow and participants might be influenced to respond in ways to conceal these biases. This effect was argued to be more present in women than men, and as such a higher number of female participants may be more influenced to answer in a socially desirable way than men. This assumption has not been well tested or verified in the research literature, however.

In an effort to counteract this disadvantage of self-report measures, a variety of procedure for evaluating so called 'implicit attitudes' have been developed and evaluated, with the most popular and easily recognisable being the Implicit Association Test, or the IAT (Greenwald, McGhee & Schwartz, 1998). Such attitudes are defined by Greenwald and Banaji (1995) as unidentified elements of past experience that "mediate favourable or unfavourable" actions or thoughts. Within the history of the participant, either through direct experience or the social context, biases have been formed that are unidentified or inaccurately discriminated by the participant and have influence on their interactions with the wider environment. Although there exists an on-going debate about the exact nature of implicit attitudes and how they influence behaviour there is evidence that scores obtained on implicit measures (such as the IAT) measure spontaneous and immediate responses, which may indicate that they are more automatic judgements than the considered and controlled behaviours typically seen on self-report measures (Freise, Hofmann, & Wanke, 2008). Research into biases based on socially sensitive attributes (such as physical appearance) may

benefit with use of these implicit measures as implicit measures do not require the participant to report their own preference or belief, as it measures their accuracy/speed in responding according to pre-specified rules and therefore the risk of a participant responding in a socially desirable 'correct' manner is reduced.

The IAT was developed based on associations within memory, in that it would be easier to associate concepts that are related in memory than to associate concepts that are not related in memory and do not have a history of similarity. The participant must respond quickly in trials that are presented as either consistent or inconsistent with pre-existing memory associations (e.g. flowers-good consistent blocks would be easier to associate in memory than flowers-bad inconsistent blocks). As such, in the IAT biases would be shown by greater response latency and greater inaccuracies in tasks that are based on concepts that are dissimilar or unrelated in memory (Greenwald et al, 1998). That is, participants on the whole will respond faster and have more correct answers on consistent trials rather than inconsistent trials. This same assumption is carried across to more socially sensitive issues and these issues are tested in the same way – in that the speed and accuracy of responses should be higher in consistent trials than inconsistent trials, and a bias would be empirically shown in the difference between latency of inconsistent versus consistent. A number of studies have been published demonstrating the IAT effect in a wide range of socially sensitive issues – homophobia, religious stereotypes and, perhaps most recognisable, racial biases (Vilathong, Lindner, & Nosek, 2010; Lemm & Banaji, 1999; Teachman & Brownell, 2001; Dunham, Baron & Banaji, 2006; Peris, Teachman & Nosek, 2008; Schmidt & Nosek, 2010). This effect has been replicated in a multitude of studies, and has been increasingly used as a measure of group based implicit biases. Despite this wealth of research, some researchers have highlighted limitations in the IAT which may reduce its appropriateness as a measure of attractiveness bias in the domain of employability.

Perhaps the most important limitation of the IAT is that it provides the relative strength of the associations among the stimuli (De Houwer, 2002) – if an IAT was used to investigate the bias towards attractiveness in the domain of employability, it would not provide independent results of attitudes towards attractive and unattractive faces. If the IAT indicated a preference for attractive individuals over unattractive individuals this could represent a positive bias to attractive individuals over unattractive individuals. Alternatively, it could indicate a bias for both but a stronger bias for attractive individuals, or it could represent a negative bias to both but a less negative bias to attractive individuals. For very nuanced investigations the researcher may require much more precise information about biases that can be provided by the IAT. Due to this limitation, a number of other implicit methods and measures have been designed that provide independent measures of the attributes being investigated, including measures such as the Go/No-Go Association Task (Nosek & Banaji, 2001) and the Function Acquisition Speed Test (O'Reilly, Roche, Ruiz, Tyndall & Gavin, 2012). These tests attempt to measure implicit preferences and biases with more detail than is found in the IAT.

A more recent development of an implicit measure to counteract self-report reliability issues in a behavioural domain (as opposed to the IAT's more cognitive or associationist domain) was the adapted version of the IAT known as the Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes et al, 2006). The IRAP draws heavily on the relational evaluation procedure (Barnes-Holmes et al, 2010), which requires participants to evaluate the stimulus relation presented. In other words, stimuli were presented with a relation (e.g. same; opposite) and asked which one was right. The IRAP expands on this by presenting relational terms that describe specific stimuli, but instead requires speed and accuracy in relations that are either consistent or inconsistent with pre-existing verbal relations (Barnes-Holmes et al, 2010). Both the relational dimension (such as same or opposite) is presented alongside a

functional dimension (such as pleasant or unpleasant); because of this an immediate relational response is produced dependent on the learning history of the individual, either verbal or nonverbal. A hypothesis similar to that of the IAT is proposed, in that response latencies for consistent pre-established verbal relations should be shorter than that for inconsistent relations. The consistent trial of flower is the same as pleasant words should have a shorter response latency than the inconsistent trial of flower is the same as unpleasant words. The IRAP has a distinct advantage over the IAT procedure, as it could be considered more direct in measuring relational responding as opposed to associations in memory which require assumptions based on indirect and immeasurable mental phenomena. Another distinct advantage of the IRAP is that the direction of the bias can be specified due to the four relational trial blocks whereas other procedures do not have this directionality built in to the underlying procedure (Murphy, MacCarthaigh & Barnes-Holmes, 2014). Due to the four trial types (e.g. attractive-employable, attractive-unemployable, unattractive employable, unattractive-unemployable) the IRAP can help determine if the bias present is pro-beauty, anti-ugly or some combination of the two. This type of nuanced responding has been absent from extant research to date, as noted by Langlois and Griffin (2006).

The IRAP involves pairs of stimuli presented to participants on a computer screen. Participants are required to respond to two types of blocks of these pairs and are required to respond accurately and as fast as possible according to two responding rules, usually two specific types of rules (Hussey, Thompson, McEnteggart, Barnes-Holmes & Barnes-Holmes, 2015). The IRAP compares, under time pressure, the ease with which participants respond according to one rule compared to the other – assessing reaction time biases that are referred to as "implicit attitudes" (De Houwer & Moors, 2010). For example, consider a hypothetical racism IRAP: two responding rules *Rule A* (Respond as if white people are good and black people are bad) and *Rule B* (Respond as if white people are bad and black people are good);

Label 1 (pictures of white people) and *Label 2* (pictures of black people); *Target 1* (loyal, trustworthy, kind, moral, generous, friendly) and *Target 2* (manipulative, dishonest, cruel, horrible, selfish, heartless); and *Response Option 1* (true) and *Response Option 2* (false), these response options are usually mapped onto the 'D' and 'K' keys. In an IRAP one of the two responding rules (Rule A or Rule B) is presented onscreen before each block of trials. In each block label stimuli are presented at the top of the screen; target stimuli are presented in the centre of the screen and response options are presented in the bottom left and right hand corners of the screen. The participant sees the responding rule screen (e.g. respond as if white people are good and black people are bad) after which clears, and the trial screen is presented (e.g. a picture of white person, the word 'loyal' and the response options 'true' and 'false'). Label stimuli contain categories, e.g. white vs. black, self vs. others, attractive vs. unattractive, and target stimuli frequently contain attributes, e.g. positive vs. negative. Both label and target stimuli contain multiple exemplars of the category or attribute.

The combination of each label categories and two target categories produce four possible trial-types (e.g. *White-Good, White-Bad, Black-Good, and Black-Bad*). The required response deemed 'correct' and 'incorrect' for each trial type is pre-determined by the task – blocks which are preceded by *Rule A* requires the participant to respond true for white faces and positive attributes and false for white faces and negative attributes. Similarly, the participant is required to respond true for black faces and negative attributes and false for black faces and positive attributes. For *Rule B* blocks, responding is reversed. The IRAP is constrained in this way to assess the difference in response times between Rule A and Rule B blocks for each of the four trial types. Participants are presented with three test block pairs of each rule, each of which contains a substantial number of trials to aid in conducting a meaningful analysis of response latency. Before these test blocks participants are presented practice blocks until they meet a mastery criterion for speed and accuracy, where the criterion

is as strict as possible to produce a meaningful effect (the IRAP effect is due to accuracy and speed pressure – getting the most correct as possible, as fast as possible).

The theoretical assumption underlying the IRAP is a behavioural approach to language and cognition – relational frame theory (Hayes, Barnes-Holmes, & Roche, 2001).

Theoretical background to the IRAP

As relational frame theory arises from behaviour analysis, the underlying theoretical background to the IRAP is distinctly different from the theoretical background from which the IAT draws from. Rather than cognitive psychology, the IRAP draws from behavioural analysis and aims to identify stimulus relations that guide behaviour and implicit attitudes.

Relational frame theory (RFT) is an approach towards human language and cognition which emphasises the importance of stimulus relations that are central to understanding language and cognition. Stimulus relations, rather than associations, allow a number of untaught stimulus relations to emerge from directly taught stimulus relations without explicit instruction or feedback. For example, simple untaught relations may arise from teaching A=B and B=C, so that symmetry may be emitted (B=A, and C=B) and transitivity (A=C). These relations may have arisen without direct training, so these relations as well as the explicitly taught relations are said to be in the same equivalence class or derived relation (Sidman, 1971). Following this stimulus equivalence, a number of research studies show a link between these equivalence classes and human language across a number of settings and contexts (Barnes-Holmes, Barnes-Holmes, Smeets, Cullinan, & Leader, 2004). An early example of the type of research into conflicting equivalence classes, which has subsequently provided the background to the development of the IRAP programme was research into religious categorisation in Northern Ireland (Watt, Keenan, Barnes, & Cairns, 1991). This research involved training for classes that were likely to interfere and be inconsistent with

pre-existing relations with the prediction that these trained classes would not be as successful or as strong as the natural relations – especially for socially sensitive topics. In the research the participants were made up of individuals who lived in Northern Ireland and English participants who did not. In Northern Ireland, due to the social context, individuals often associate particular names and symbols with either the Protestant or Catholic region or religions but this practice is relatively unheard of in England. Training involved matching Catholic names with nonsense syllables and these syllables with Protestant symbols, following this training participants were given an equivalence test where they were asked to match the names directly with the symbols (so that the taught A=B and B=C relations would encompass the untaught A=C relation). Many of the Northern Irish participants could not complete this test, but the English participants could. It appeared that the pre-existing relations established within in the social context of Northern Ireland were too strong to allow the formation of the new inconsistent relations, while the English participants had no such pre-existing relations within the social context and could form the new relations with relative ease. These effects have been replicated among a number of other domains, including terrorism (Dixon, Rehfeldt, Zlomke, & Robinson, 2006), self- esteem (Merwin & Wilson, 2005) and self-concept (Barnes, Lawlor, Smeets, & Roche, 1996). This approach pitting already pre-existing relations against conflicting laboratory created relations provides the conceptual framework for the IRAP (Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010).

In this case, unlike the IAT, responses on the IRAP are driven by relations rather than associations. As a behavioural account of stereotype bias, the IRAP can avoid appealing to internal explanations such as memory associations by employing the Relational Elaboration and Coherence (REC) model. In this model the IRAP trials produce an immediate relational response (brief immediate relational responding, BIRR). This response is often determined by

the history of the participant and the current contextual variables, and this immediate response is the one that will be emitted first most often under a time constraint. In consistent trials that match this immediate relation the response will be emitted quite quickly, in contrast inconsistent trials that do not match this immediate relation will require additional time to offset the automatic relation in order to respond accurately and so will be emitted less quickly. As the IRAP requires immediate relational responding, the latency between presenting the trial and the response will provide information on the immediate relation brought about by the participant and as a result, a bias that may be present (Barnes-Holmes et al, 2010). This also explains the differences in bias and bias strength between the IRAP and explicit measures, when completing explicit measures such as questionnaires there is no time limit that prevents additional relational responses that cohere with a wider relational network. For example, when asked to rate pictures of black and white men as dangerous or safe the immediate relational response might be black = dangerous, but there would be additional relational networks such as "both men are similar", "both men look professional" or "discrimination based on race is wrong". This relatively complex relational network is easier to engage with in the absence of a timed trial, whereas the IRAP requires an immediate response and so draws from the immediate relational response – or the automatic response as opposed to the carefully considered response. This is shown quite successfully when reducing the time latency from 3,000 to 2,000 ms and the corresponding increase in racial stereotyping effects (Barnes-Holmes, Murphy et al, 2010). This finding supports the argument of an immediate and automatic response that often displays a form of bias that is mediated by additional relations in a non-timed response trial.

The IRAP has been used in more socially sensitive contexts than the previously mentioned study with similar effects: smoking attitudes (Vahey, Boles & Barnes-Holmes, 2010); ageist attitudes (Cullen, Barnes-Holmes, Barnes-Holmes & Stewart, 2009); self-

esteem (Vahey, Barnes-Holmes, Barnes-Holmes & Stewart, 2009); body weight biases (Rodd, Stewart & Barnes-Holmes, 2010); and the implicit beliefs of sexual offenders (Dawson, Barnes-Holmes, Gresswell, Hart, & Gore, 2009). In these studies, a number of them reported evidence of implicit biases found with the IRAP that are not present in explicit measures measuring the same domain. As well, there exists a body of research that indicates findings of the IRAP are predictive of behaviour (such as Carpenter et al, 2010 – where results indicating a positive cocaine bias on an IRAP pre-treatment was associated with lower attendance rates and proportions of cocaine free urine in a treatment program).

Further research surrounding the IRAP has indicated that the IRAP is not amenable to 'fake' or controlled responding, even when instructed to control their responding (McKenna, Barnes-Holmes, Barnes-Holmes & Stewart, 2007). Participants were presented a standard IRAP program. Once this was completed the participants were informed about how the IRAP works and half the participants were asked to try to fake the next IRAP, while the other half were explicitly told how to fake the IRAP. Results found no evidence of faking. This ostensible insusceptibility of the IRAP to faking responses, even when explicitly told how to fake responses, is an advantage of the measure over the explicit measures currently employed in employability research.

The IRAP and the Use of Pictorial Stimuli

Physical appearance and employability is a relatively recent area of research for the IRAP and as such the stimuli used in the IRAP should be appropriate for assessing attitudes towards attractiveness or unattractiveness. Research using the IRAP has employed both pictures (such as in an IRAP designed to measure racial biases) as well as words (such as in an IRAP designed to measure attitudes towards those with a mental illness). The IRAP programme itself is amenable to using either word or picture stimuli and can be easily

adapted to the types of stimuli deemed suitable for the relevant attitude or bias in question. It has been well established that pictures have been found to produce effects and responses that were stronger or exaggerated than found with words. Pictures have been found to be more effective when presenting information and when processing information (Bourisseau, Davis, & Yamamoto, 1965). It is argued that pictures are easier to process than words, and as such might arouse a reaction more directly and much speedier than words (Dasgupta, McGhee, Greenwald, & Banaji, 2000; Smith, Bradley, & Lang, 2005).

This research suggests that using pictures within an IRAP may provide a more accurate measure of the immediate response of an individual due to the emotional reactions evoked by pictures rather than word descriptions. For the measurement of biases directed towards physical attractiveness, the use of pictures as a target stimulus is doubly important as the use of word descriptions would interfere with any particular response that would be emitted towards the pictorial representations of physically attractive and physically unattractive individuals.

Although there is a wide bank of research that supports attractiveness bias in both real life circumstances outside of experimental conditions and in experimental conditions, there is a limited history of researching attractiveness bias using implicit measures such as the IAT or the IRAP (Murphy, MacCarthaigh & Barnes-Holmes, 2014; Murphy, Hussey, Barnes-Holmes & Kelly, 2015). As physical attractiveness appears to affect many aspects of the domain of employment (such as employability, income, and promotion in the workplace) it remains an important area of study within psychology. Previously in physical attractiveness research the majority of studies have been conducted with explicit, self-report measures which are derived primarily from the social psychology theoretical background, whereas the IRAP allows a behavioural investigation into this domain that may allow more nuance and detailed investigation into the bias. As such a socially sensitive area, biases towards physical

appearance are susceptible to socially desirable responding as the attitudes of an individual towards physically attractive individuals may differ from the societal image of attractive individuals (i.e. while a participant might be biased towards attractive individual, the socially desirable response might be no bias towards attractive or unattractive individuals). Therefore, the explicit measures used in research surrounding physical attractiveness and employability may generate results that are influenced by the views of the social context in which in the individual resides. Further, responding on self-report measures can also be influenced by introspection problems (Murphy et al, 2014). Participants may not be aware of biases they might hold towards certain social groups and self-report measures can only evaluate the extent to which participants are aware of their own bias. The IRAP, however, is a measure which can measure responses towards physical attractiveness and employability that appear to be independent of the social desirability that appears to be affecting results on the explicit measures (Roddy, Stewart, & Barnes-Holmes, 2010).

In their study, Murphy et al used the IRAP and a number of explicit measures in order to investigate the strength and direction of attractiveness bias. Using the consistent block of attractive-successful and the inconsistent block of unattractive-successful, the study interpreted the response latency in order to ascertain both the existence of an attractiveness bias and the directionality of any such bias. Response latencies indicated that there existed significant pro-attractive biases as well as significant anti-unattractive biases. This implicit measure moderately correlated with the implicit measures used in the study, as well as results from the previous attractiveness bias literature which indicated a 'beautiful is good' direction often shown in research (Shahani-Denning, 2003), as well as an often-overlooked dimension that 'unattractive is bad' that appears to be a relatively under looked aspect of the attractiveness bias. Murphy, Hussey, Barnes-Holmes and Kelly (2015) extended this research by again measuring attractiveness bias for successfulness again – this time more focused on

directionality. There again exists a bias for both male and female participants, although the bias was shown to be pro-attractiveness and not anti-unattractive. Additional research is needed to confirm this directionality, both in implicit measures and explicit measures. An important aspect of this study also investigates gender effects for participants, and although there was a bias for both genders, there was not a significant difference due to participant gender, however the gender of target stimuli pictures may also play a part in the strength of the bias (Heliman & Saruwatari, 1979; Nolan, Murphy & Barnes-Holmes, 2013). The current research aims to use the IRAP to approach attractiveness bias with a more nuanced methodology than previous research in the area, and aims to examine gender effects of participants, as well as investigating the directionality (pro-beauty, anti-unattractive, or combination of both) and strength of an attractiveness bias.

Although research has been conducted into investigating attractiveness bias in terms of self-report measures (Dion, Berscheid & Walster, 1972) as well as discrimination tasks (Cash, Gillen & Burns, 1977), there has not been an evaluation of implicit measures and their effects on actual behavior in the domain of biases towards physical attractiveness. This investigation is important as the main goal of implicit measures should be to predict real-life behaviour (De Houwer, 2002). Similarly, there has not been an investigation into the malleability of these biases and the extent to which they are changeable. Roddy, Stewart and Barnes-Holmes (2011) investigated the extent to which the IRAP could predict explicit discriminatory behaviour in regards to a pro-thin, anti-fat bias. Surprisingly, there was a significant correlation with pro-slim, anti-fat bias on the D-IRAP scores and participants reporting that they had increased likelihood of interacting with an overweight individual. This may be a result of social desirability effects on the explicit discriminatory measure – especially as the measure was solely dependent on weight, and not on other information such as a CV (as in Vanman et al, 2004).

The IRAP has predictive power for the behaviour of individuals as seen in Carpenter, Martinez, Vadhan, Barnes-Holmes and Nunes (2010). In this study, an IRAP was conducted to investigate beliefs surrounding cocaine use and the correlations between results on the IRAP and treatment outcomes. The study found that stronger beliefs about the positive effects of cocaine use prior to treatment were associated with poorer treatment outcome. However, the same association was not found for beliefs measured with self-report measures. The current research programme aims to investigate this link between scores on the implicit measures and explicit discriminatory behaviour when asked to complete a behavioural task (a forced CV choice between attractive and unattractive individuals) in relation to physical attractiveness bias.

Similarly, the IRAP has successfully predicted an avoidance of live spiders based on IRAP scores (Nicholson & Barnes-Holmes, 2012). The IRAP predicted the avoidance of live spiders but also reliably predicted scores on a widely used explicit measures for fear of spiders and it could distinguish between groups rated as high or low fear. These studies provide the rationale for investigating the predictive power of the IRAP for choosing Curriculum Vitaes based on appearance – if participants reliably choose CVs with attractive pictures attached over CVs with unattractive pictures attached and scores on the IRAP can reliably predict this, then the IRAP could be a useful clinical tool to screen potential interviewers for a bias towards attractive individuals over unattractive individuals.

The current research programme aimed to investigate physical attractiveness bias in terms of employability, by a) developing an IRAP that specifically targeted physical attractiveness and employability for both attractive and unattractive individuals and b) employing a number of explicit, self-report measures in the domain of employability (the Measures of Interpersonal Attraction and the Interpersonal Judgement Scale (McCroskey and McCain, 1974; Byrne, 1971)). This research is the first attempt to use the IRAP in an

investigation of attractiveness bias towards employability (previous research has investigated attractiveness bias in the context of evaluations of general 'successfulness'; Murphy, MacCartaigh & Barnes-Holmes, 2014) as well as an investigation into the effects of gender on ratings of employability. The first study in the research programme investigated both the strength and directionality of the bias using the IRAP, as well as the extent to which the results of the IRAP correlate to explicit measures of attractiveness bias (as in Murphy, MacCarthaigh & Barnes-Holmes, 2014), while the second study in the research programme investigated the gender effects of stimuli in the IRAP towards employability for attractive and unattractive women, and attractive and unattractive men. This research may highlight biases that remain undetected in self-report measures due to the IRAP four trial type methodology which can be used to conduct an in-depth investigation.

Study 1. The first study consisted of a picture based IRAP, which was designed to measure attitudes towards attractive and unattractive faces in the context of words relating to good and bad employee traits. Explicit measures (the Measures of Interpersonal Attraction scale and the Interpersonal Judgement Scale) were also employed for comparison as well as informative purposes. Additionally, the extent to which IRAP scores predict actual behaviour, was investigated using a discriminatory behavioural task (a forced CV choice: similar to that shown in weight bias in Roddy, Stewart and Barnes-Holmes, 2011).

Study 2. Based on the results of study 1 and the research indicating the need for more nuanced investigation into the level of attractiveness and pro-beauty bias, an IRAP was developed using stimuli of high and medium attractiveness and the same label stimuli as study 1. Study 2 used the same explicit measures and behavioural task as study 1.

Study 3. The final study aimed to investigate the presence of a 'beauty is beastly' effect for female stimuli and traditionally 'male', executive job titles. An IRAP was designed

which consisted of male and female images, and six traditionally male, executive jobs and six traditionally female, nonexecutive jobs. Explicit measures were also employed for comparison as well as informative purposes.

Chapter 2

Study 1

Using the IRAP to Measure Implicit Pro-Attractive Stereotype in Evaluations of Employability

The first study in the current research thesis aimed to assess participant bias towards physical attractiveness and employability using 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) and a behavioural implicit measure known as the Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes et al, 2006). The current study is the first attempt to use implicit measurement towards physical attractiveness and employability, as previous research in the area has mostly employed explicit self-report questionnaires measures (e.g. Griffin and Langlois, 2006) or behavioural task such as Curriculum Vita (CV) selection (e.g. Heilman & Saruwatari, 1979), and previous research in physical attractiveness using the IRAP has focused primarily on physical appearance and evaluations of successfulness, but not employability (e.g. Murphy, MacCartaigh et al, 2014; Murphy, Hussey, Barnes-Holmes & Kelly, 2015). As previous research has indicated that university students' ratings of job applicants were nearly identical to those of professionals (Gordon, Slade & Schmitt, 1986), as well as the fact that college students may be representative of young middle-class adults who may form part of interview panels (Murphy et al, 2015) university students were selected as the research participants for the current study.

Thus, the current research aims to extend the implicit research literature on beauty stereotype in the context of employability using the IRAP four trial type methodology to provide a more nuanced investigation into the directionality of the bias (e.g. pro-beauty, anti-ugly, or both), while also extending the IRAP research literature in this domain. The IRAP developed in the current study utilised both word and picture-based stimuli. The latter were photo images of facial composites that were attractive and unattractive, six of each. The word-stimuli were six positive employee attributes (intelligent, trust-worthy, hard-working, reliable, approachable, and responsible) and six negative employee attributes (stupid,

dishonest, lazy, unreliable, standoffish, and irresponsible) were used in the IRAP programme. The computer-screen presented a sample stimulus (either an attractive of unattractive facial image), with an evaluate stimulus (positive or negative attribute), and the response options 'true' or 'false'. These relations were presented in blocks of trials, which were deemed either *Rule A* (Consistent with stereotype, e.g., attractive-intelligent-true/unattractive-intelligent-false) or *Rule B* (*Inconsistent with stereotype, e.g., unattractive-trustworthy-true/attractive-trustworthy-false*). In essence, four trial types were involved: (i) *Attractive-Employable,* (ii) *Attractive-Unemployable*, (iii) *Unattractive-Employable*, and (iv) *Unattractive-Unemployable*. University students were selected as research participants for the current study, not only for the usual reason that they can function as a convenient, accessible participant sample, but also because previous research has indicated that university students' ratings of job applicants were nearly identical to those of professionals (Gordon, Slade & Schmitt, 1986).

The predictions for the current study included: (1) the IRAP will indicate a bias towards attractive individuals as employable and unattractive individuals as unemployable; (2) that explicit measures will indicate a bias towards attractive individuals as more employable; and (3) participants will select the CVs of attractive individuals over the CVs of unattractive individuals. These predictions are based on the research conducted in physical attractiveness and employability, in that results across multiple studies have shown the presence of a pro-attractive bias toward employability using self-report measures (Shahani-Denning, 2003). In addition to this, research has also shown the presence of pro-attractive-employable bias in CV selection (Cash & Kilcullen, 1985). As there is such evidence of a pro-attractive-employable bias in explicit measures, implicit measures similarly showed participant bias in favour of attractiveness bias in relation to attractiveness and successfulness using the IRAP and attractive facial images (Murphy, MacCartaigh & Barnes-Holmes, 2014).

Method

Participant details

Twenty-four individuals completed the study; six others participated but their data were discarded as they did not meet the IRAP inclusion criteria explained in the results section. The final study included 24 university students comprised of 11 males and 13 females ranging in age from 18-28, mean age 21.81. All participants completed an IRAP and four explicit questionnaires and rating scales. Participants were recruited from first year undergraduate students attending Maynooth University.

Ethical Considerations

The IRAP procedure is a computer programme which is considered to pose a negligible risk in the current study, however participants were advised that if they had a history of seizures they should excuse themselves as participants. Before beginning the experiment, participants were provided with an information sheet and a consent form and were asked to read the information sheet and sign the consent form if they understood the experiment and wished to participate. Approximately half way through the experiment procedure, participants were again reminded that participation was voluntary, and that they were free to withdraw at any stage during the experiment procedure without penalty. Participants were informed that no identifying information about themselves would be kept, and that they would be referred to only by code numbers once they finished the experiment. They were also informed that due to the lack of identifying information, once they had completed the experiment the experimenter would not be able to identify their data individual and as such they would not be able to remove their data from the study once the experiment session had completed and they had left. Participants were informed that data were analysed at group level and not at an individual level, and only group results could be provided if they

wished to see the data. In addition, participants were informed that data would be stored in an encrypted file and retained for the length of time specified by university and legal requirements.

Apparatus and Materials

The research was conducted in Maynooth University, in the psychology departments testing labs. The IRAP (2012 update version) was presented on a standard laptop computer (a Toshiba Satellite, Intel core i3). The IRAP program controlled all stimulus presentation and recording of participant output (both participant time latency and accuracy). The stimuli presented by the IRAP were made up of six "good employee" words (e.g. responsible, intelligent) and six "bad employee" words (e.g. irresponsible, stupid). These words were developed from a focus group (N=5, 2 males and three females, mean age 21.8) containing members independent of the participants who completed the IRAP. The IRAP also presented twelve pictures comprising of six male faces (three attractive and three unattractive) and six female faces (three attractive and three unattractive). These pictures were independently rated on attractiveness by the same participants in the word stimuli focus group (see Appendix 1). In addition, the words "true" and "false" were presented as response options. The stimulus arrangements used in the study are presented in Table 1.

Before commencing, each participant's age and sex were recorded by the experimenter. A series of rating scales and questionnaires were presented to the participant. These included the Interpersonal Judgement Scale (IJS) and Measures of Interpersonal Attraction scale (MIA; see Appendices 2 and 3).

The MIA consists of three separate scales relating to social attraction, physical attraction and task attraction. For the purposes of this study, only the task attraction subscale was presented to participants. This scale consists of six statements with a 5-point scale before

Table 1: *Stimulus arrangements presented by the IRAP*

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

the presentation of the scale (1 = strongly disagree and 5 = strongly agree). Participants were asked to complete this scale for each picture of attractive and unattractive individuals by writing the number which best describes their agreement/disagreement with each statement (e.g. I could not get any work done with him/her). The IJS consists of four 'distractor' questions presented before the two questions which are the measure of interpersonal task attraction. Participants were required to complete this scale for each picture by writing the number which best describes their agreement/disagreement with each question using the same 5-point scale (1 = strongly disagree and 5 = strongly agree).

In additions, two Likert scales were presented for each picture (see appendix 4). These scales presented the words '*attractive'* and '*employable'* and participants scored each picture by circling a number from -3 (very unattractive/unemployable) to +3 (very attractive/employable). These measures were designed to provide an explicit measure of participants' views on the attractiveness and employability of attractive and unattractive faces. Finally, a brief behavioural task was employed to ascertain if behaviour was predicted by these measures. This task involved presenting forced choices of CVs for a human resource manager job vacancy between CVs with attractive pictures attached and CVs with

unattractive pictures attached. There were four pairs – with four attractive CVs and four unattractive CVs overall, with CVs having similar qualifications and work experience and the only difference being the picture attached to the CV.

Procedure

Participants were first briefed as to the nature of the study and were presented with an information sheet giving details of the research (see Appendix 5) and were required to read and sign a consent form (see Appendix 6). Once consent was received, each participant was presented with the four explicit measures which included the two rating scales and two questionnaires. Following these questionnaires and rating scales, they were presented with the CV choice task. Participants were instructed that the CVs were hypothetical finalists for a vacancy in the human resources department in a company, and that they had to choose one CV from each pair given to them. Each pair of CVs contained one attractive and one unattractive CV, and four pairs were presented to the participants, each pair of the same gender. Once these explicit measures had been completed, the participant then was presented with the IRAP task.

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. On each trial, either a ‘good’ word or a ‘bad’ word would appear in the centre of the screen along with a picture (either an ‘attractive’ or ‘unattractive’ picture) at the top of the screen. In addition, the response options ‘true’ and ‘false’ would be presented in the bottom left and right hand side of the screen respectively. Participants were informed that the response options were represented 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 80% correct responding with a response latency below 2000ms. An example of each trial type was provided to the

participant, and participants were told that there were practice blocks provided until the aforementioned criterion was met. They were also informed that after each block, the instructions and correct answers would be reversed, requiring the opposite pattern of responding from the previous block. Finally, participants were told that in some parts of the experiment they may be asked to respond in a way that did not reflect their beliefs but that this was part of the experiment. They were also told that the researcher would sit with them for the first practice block to help illustrate the task and the response requirements.

Each trial of the IRAP involved presenting of the sample pictures (six 'attractive' and 'unattractive' faces) at the top of the screen, with one of the label stimuli (six 'good' words and six 'bad' words) in the middle of the screen. As well, the two response options 'true' and 'false' appeared in the bottom left and right hand corners respectively. 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. If the response option chosen was the option deemed correct for that trial, the screen cleared for a 400ms interval, and the next trial was presented. If a participant chose the response option deemed incorrect (including pressing the wrong key), then 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 interval. 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 pro-attractive/anti-unattractive for good words, and pro-unattractive/anti-attractive for bad words i.e. participants were required to respond 'true' when an attractive face was presented with a good word or unattractive face was presented with a bad word, but were required to respond 'false' when an attractive face was presented with a bad word or an unattractive face was presented with a

good word; see figure 1. When the first 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/anti-attractive for good words and pro-attractive/anti-unattractive for bad words, i.e. participants were required to respond 'false' when an attractive face was presented with a good word or an unattractive face was presented with a bad word, but were required to respond 'true' when an attractive face was presented with a bad word or an unattractive face was presented with a good word. When both blocks were completed the 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 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 2000ms) 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. There were no criteria required for these blocks, and therefore once each block was completed the participants continued until all six blocks were finished. Responding with 'consistent' pro-attractive for good words and pro-

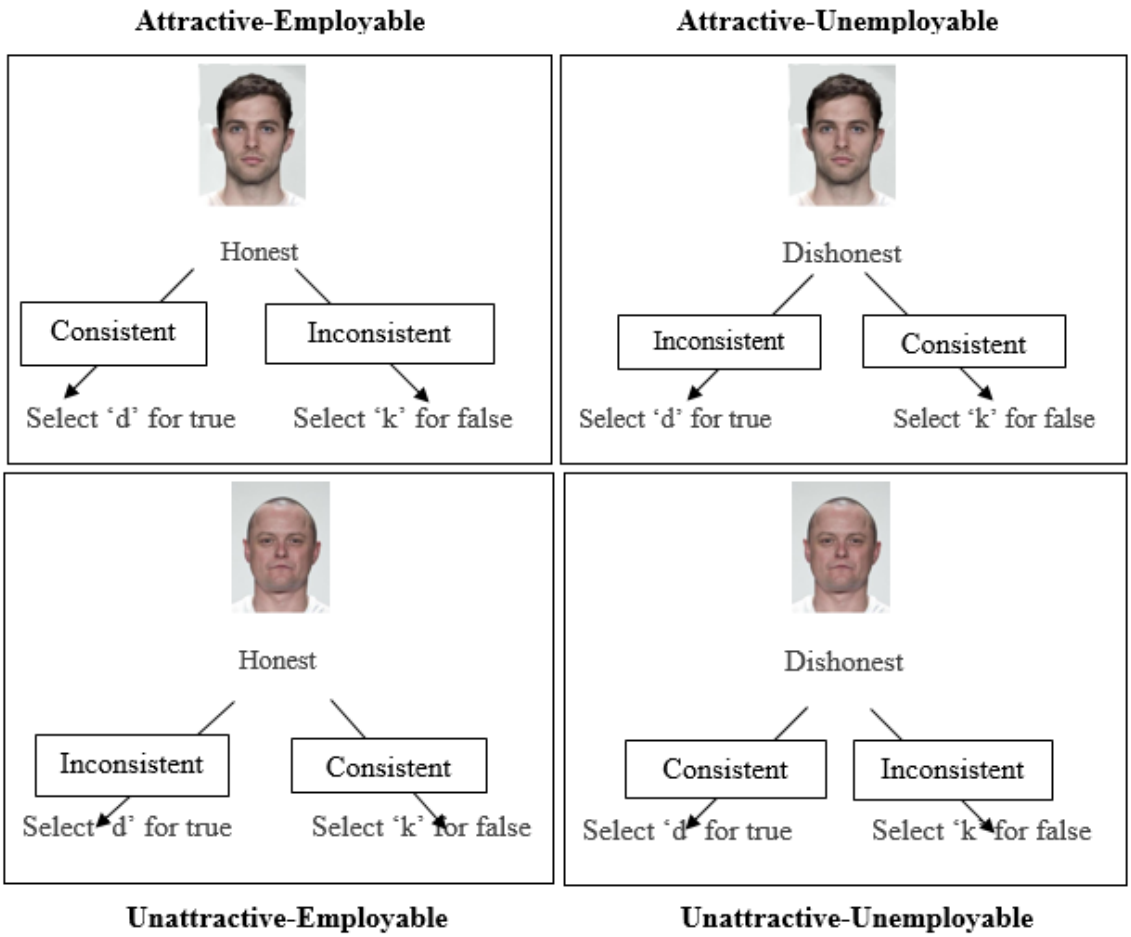


Figure 1. Representations of four IRAP trial-types. The superimposed arrows and labels indicate what would be considered a bias towards attractive-employable (consistent) or a bias towards unattractive-employable (inconsistent) response for each trial type. (These are for illustration purposes and did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

unattractive for bad words and responding with 'inconsistent' pro-unattractive for good words and pro-attractive for bad words was required on alternating blocks. Once test blocks were completed, a message appeared informing participants that this part of the experiment was completed and to contact the researcher. Upon completion of the study, the participants were debriefed and thanked for their participation.

Results

Overview of data analysis

Analysis of the data from the explicit measures is presented first followed by analysis of the IRAP data. The analysis includes comparisons of means for the attractive and unattractive explicit measures; for IRAP data one sample t-tests, correlational analyses, analysis of variance tests (ANOVAs) and regression analysis were conducted. One sample t-tests were conducted to examine individual IRAP effects, correlational analysis were conducted to examine relationships between the IRAP and explicit measures and an ANOVA was conducted to examine the effects of gender of participants on IRAP scores for each of the four trial types. Finally, a regression analysis was conducted to examine the predictive power of the IRAP for a CV selection task (selecting between attractive vs. unattractive individuals).

Explicit measures

In order to conduct analysis of the data and consistent with the scoring guide for each explicit measures scores for the task attraction subscale of the Measures of Interpersonal Attraction (MIA) were calculated for each participant for both the attractive and unattractive images. This scoring yields two MIA scores for each participant; similarly scores for the Interpersonal Judgement Scale (IJS) were calculated for each participant for both attractive and unattractive images. Four scores for each participant were produced from the two scales (one for attractive images and one for unattractive images). The means and standard deviations for each of the measures and conditions are presented in table 2. For each of the two Likert scales scores were calculated for each participant for both attractive and unattractive images, yielding a further two scores for each participant. The means and standard deviations for each of the Likert scales and conditions are also presented in table 2. To examine if there was a statistically significant difference in scores on each of the four

measures for attractive and unattractive images a number of Wilcoxon signed rank tests (for non-normally distributed scores) were conducted. The results of these tests are also summarised in table 2. As predicted, attractive images were rated as statistically significantly more attractive than unattractive images and were rated as statistically significantly more employable than unattractive images across the three employability measures. For the attraction Likert scale a Wilcoxon signed rank test indicated that there was a statistically significant difference in scores for attractive and unattractive faces, $z = -4.29$, $p < .001$. For the employability Likert scale a similar Wilcoxon signed rank test indicated that there was a statistically significant difference in scores for attractive and unattractive faces, $z = -3.79$, $p < .001$. 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. For all scales there was statistically significantly higher score for attractive images over unattractive images, which indicates that attractive images were scored as more attractive and more employable than unattractive faces. Effects sizes ranged from .50 to .71 which indicates a strong effect size using Cohen's (1988) criteria. These analyses therefore confirmed that the variables utilised in the study were functioning as planned.

In addition to each of the explicit measures completed, a behaviour 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) an equal amount of participants selected majority attractive CVs as were neutral and selected an equal number of attractive and unattractive CVs. Only 17% of participants selected all attractive CVs and 25% of participants selected three attractive CVs (out of a possible four). 41% of participants were neutral, selecting an equal number of attractive and unattractive CVs, indicating no

preference for attractive faces. Only 17% of participants selected three or more unattractive CVs over attractive CVs. In summary, there was no consistent selection of attractive CVs over unattractive CVs.

Table 2: *Summary of Descriptive and Inferential Statistics for the Explicit Measures*

Explicit Measures	**Attractive Faces**		**Unattractive Faces**			
	M	*SD*	*M*	*SD*	*p score*	*Effect Size*
MIA	132.13	16.34	105.96	18.82	.001	.50
IJS	43.09	5.48	30.74	4.27	.000	.59
Attraction Likert Scale	13.17	3.52	-5.42	6.76	.000	.71
Employability Likert Scale	10.79	5.98	.50	8.23	.000	.54

IRAP data

The primary data used for analysis were *D*-IRAP scores obtained from an adaption of the *D* algorithm used in the IAT (Greenwald et al, 2003). Using response latency (the time in milliseconds between the onset of the trial and a correct response) the data for each participant were converted into *D*-IRAP scores (as in Hussey, Thompson, McEntteggart, Barnes-Holmes & Barnes-Holmes, 2015). The steps involved in calculating the *D*-IRAP scores were as follows: (i) only response-latency data from the test blocks were used; (ii) latencies above 10,000ms were removed from the data set; (iii) if more than 10% of the test block trials had latencies less than 300ms the data for that participant were removed; (iv) twelve standard deviations for the four trial types were calculated: four for the response latencies for test blocks one and two, four from the latencies of test blocks three and four, and

four from the latencies from the test blocks five and six; (v) twenty-four mean latencies for the four trial types in each test block were calculated; (vi) difference scores were calculated for the four trial types, for each pair of test blocks, by subtracting the mean latency of the pro-attractive block from the mean latency of the corresponding pro-unattractive block, (vii) each difference score was then divided by its corresponding standard deviation (step iv), yielding twelve *D*-IRAP scores, one for each trial type for each pair of test blocks; (viii) four overall trial type *D*-IRAP scores were calculated by averaging the three scores for each trial type across the three test blocks; (ix) an overall *D*-IRAP score was calculated by averaging all four trial type *D*-IRAP scores from step viii.

This transformation yielded an overall *D*-score that was positive, indicating that overall there was a bias towards Rule A over Rule B (i.e. faster and more accurate responding to "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 (a perfectly neutral response), t (23) = 3.88, $p < .001$, eta square = 0.25. This transformation also yielded four *D*-scores, one for each trial type. A positive *D*-score indicates a pro-attractive employable bias and negative scores indicate a pro-unattractive employable bias. In order to compare directly across the four trial types, and for the purposes of further data analysis using SPSS, the signs for trial type three and four were inverted (e.g. reverse plus scores to minus, and minus scores to plus; see full explanatory details in Hussey & Barnes-Holmes, 2013). This inversion was to facilitate comparisons across trial types and does not alter the absolute value of IRAP *D*-scores. Following this additional inversion, positive *D*-scores indicate a pro-beauty and pro-unattractive bias, and negative *D*-scores indicate an anti-beauty and anti-unattractive bias. All other presentation and graphs of IRAP *D*-scores in the current thesis are presented using uninverted scores for ease of understanding. On the graph, scores above 0 indicate a pro-employable bias towards attractive images (i.e.

'pro-beauty' and 'anti-ugly') and scores below 0 indicate a pro-employable bias towards unattractive images (i.e. 'anti-beauty' and 'pro-ugly').

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 (see below). 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 was biases toward attractive images as employable and not unemployable, but no bias towards unattractive images in either direction.

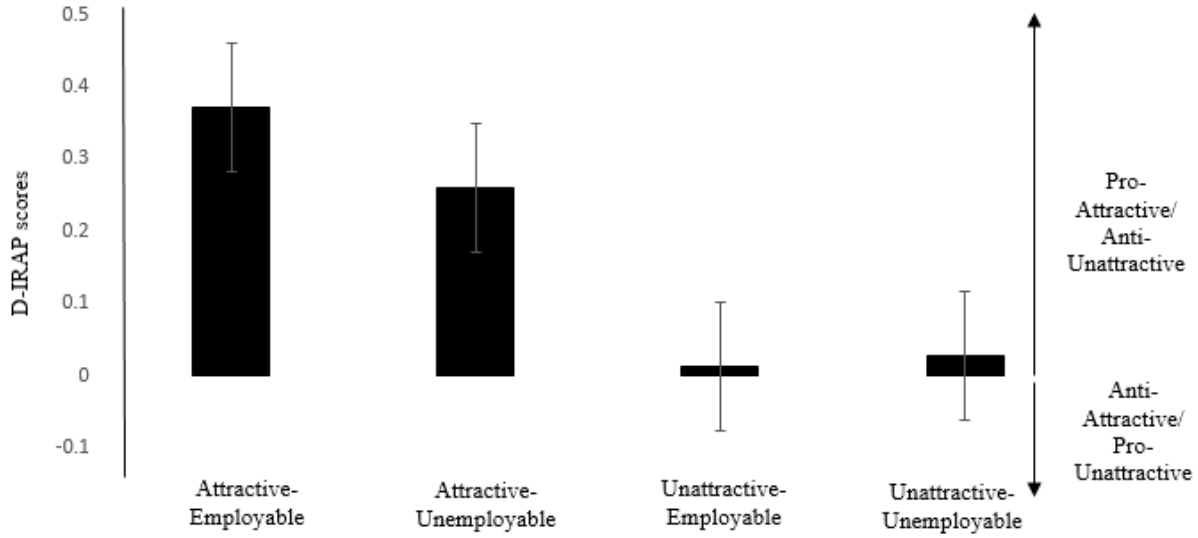


Figure 2: Mean *D*-IRAP scores for each trial type.

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/Trial type 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 = .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 effect on trial type scores and there was no overall differences between males and females. There was a statistically significant main effect for trial-type (Wilks Lambda = .34, F (3, 20) = 30.63, $p < .05$). Post hoc comparisons using a number of paired-sample t-test indicating that only scores between *Attractive-Employable* and *Unattractive-Employable*, and *Attractive-Employable* and *Unattractive-Unemployable* trial types were statistically significantly different from each other at the Bonferroni alpha level ($p < .008$).

Correlations between implicit and explicit measures

Each of the *D*-IRAP scores for the four trial types were entered into a correlation matrix with the four explicit measures (eight scores, four for attractive images and 4 for unattractive images. Of the resulting correlations, three were statistically significant and these are presented in table 3. Of the four trial types, only the *Attractive-Unemployable* and *Unattractive-Unemployable* trial types correlated with the explicit measures, one related to attractive images and two related to unattractive images.

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 as an implicit bias

towards unattractive images as unemployable increased scores on the MIA decreased for unattractive images.

Prediction of CV choice

In order to determine if the IRAP trial types had predictive effects on a behavioural 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 non-significant 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 (the IJS and the MIA) as well as on the Likert rating scale for employability. For the behaviour task (a CV choice task based on attractive/unattractive individuals), there was an equal amount of participants who chose the majority of attractive CVs and participants who chose equal amount 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 images as not-unemployable across two trial types was indicated. Across the four trial types, the *Attractive-Employable*

Table 3: *Correlations between IRAP trial-types and explicit measures.*

IRAP trial type	**Attractiveness Rating**		**Employability Rating**		**MIA**		**IJS**	
	Attractive	Unattractive	Attractive	Unattractive	Attractive	Unattractive	Attractive	Unattractive
Attractive-Employable	.15	.24	.05	-.10	.18	-.19	.11	-.11
Attractive-Unemployable	.41*	.76	.11	.16	-.03	.14	.78	.43*
Unattractive-Employable	.15	.23	-.27	.02	.01	.03	-.25	-.06
Unattractive-Unemployable	.01	-.11	.08	-.31	.33	-.47*	.10	-.30

*$p < 0.05$

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.

Discussion

The current study 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). An IRAP was developed to measure participants' implicit evaluations of the employability of individuals either rated as attractive or unattractive. Results from the IRAP were then compared with results from the two explicit scales in order to explore differences in responding across both explicit and implicit measures.

Overall, 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. As predicted, the Likert rating scale for attractiveness was statistically significantly higher for individuals with attractive faces than

unattractive faces. These results mirror results found in previous self-report studies conducted on attractiveness and employability. For the behaviour task (a CV choice task based on attractive/unattractive individuals), only a small majority of participants choose three or more unattractive CVs (out of a possible four) – 56.2% compared to 48.8% of participants who choose a majority of attractive CVs. This result was unexpected, as previous studies indicated that attractive CVs were chosen consistently over unattractive CVs. One possible limitation is that these CVs were presented following the explicit measures and as such participants' responses may be due to an element of socially desirable responding.

The overall *D*-IRAP score indicated that there was a statistically significant pro-attractive bias in participant responding, that is that responding according to the consistent *Rule A* was faster and more accurate than responding according to the inconsistent *Rule B*. 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 regards to attractive images but neutral to unattractive images. 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 et al, 1977). However, this finding was not consistent with the gender effects of stimuli as found in the beauty is beastly studies (Heilman and Saruwatari, 1979) as well as when evaluators were the same sex as stimuli.

Correlational analyses revealed that there were three statistically significant correlations between the four IRAP trial types and the explicit measures. 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 none of the IRAP scores, both *D*-scores from the four trial types and the overall *D*-scores, were non-significant predictors of CV choice across the participants, with none of the regression models reaching a statistical significance level. However, this may be explained by the presentation effects of the CV choice task, as they were presented after the explicit measures 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 behavioural task then it may limit its utility in measuring the beauty-bias effect.

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 also allowed a more nuanced insight into the directionality of the bias. Critically, it revealed the bias is in the direction of attractive individuals are more employable and not that unattractive individuals are more unemployable. This supports the validity of using the IRAP to investigate attractiveness and may, in the future, allow for the development of intervention materials that target the specific bias present. Nonetheless, a criticism of the current study, and indeed many studies investigating physical attractiveness bias, is that the current study only utilises examples of individuals of high and low attractiveness and not examples of individuals of medium attractiveness (Langlois et al, 2000). Due to the fact that people are more likely to encounter people of middling physical attractiveness than people of very low or very high physical attractiveness in real world settings, it is possible that an IRAP consisting of more ambiguous images may result in different IRAP effects. The aim of the second study was to determine if

presenting pictures of individuals of medium attractiveness would produce different effects to those found in study 1. This could extend the implicit responding literature by providing yet more nuanced data and support a more ‘ecological’ investigation of the beauty bias effect.

Chapter 3

Study 2

Using the IRAP to Measure Implicit Pro-Attractive Stereotype in Evaluations of Employability Using Images of Medium Physical Attractiveness

Participant bias in a cohort of twenty-four undergraduate university students towards physical attractiveness and employability ratings was assessed using both implicit (the IRAP) and explicit (self-report questionnaire) measurements in the first study of the current research. Findings in study 1 indicated that participants implicit measures indicated that a pro beauty bias was evident, whereas there was no statistically significant bias in either direction for unattractive individuals. Participants consistently rated attractive images as more employable on explicit measures and there was no effect of gender on scores on the IRAP and IRAP scores had no predictive power on a CV selection task.

The second study of the current research aimed to determine if the findings of study 1 could be replicated using 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, as little research has been conducted using more ambiguous images of stimuli of medium attractiveness (Langlois et al, 2000; Griffin and Langlois, 2006). Specifically, on the same measures used in study 1 would participants rate attractive individuals as more employable than those of medium attractiveness, and a further question was whether the same directionality of the bias towards attractive images be found (i.e. pro-attractive rather than anti-unattractive).

The design of study 2 was identical to that of study 1, except the images used in the IRAP and explicit measures. The images of unattractive individuals were replaced with images of medium attractive individuals, but label stimuli and the attractive images remained the same. Specifically, the same stimuli as per Study 1 were used, but the unattractive facial image stimuli were modified or "softened" to facilitate a more nuanced investigation; specifically, is a pro-beauty bias shown even when the differences between attractive/unattractive facial images is more subtle. Such an approach may have ecological

validity, in that potential employees may not commonly fall into generally agreed categories of attractiveness. Study 2 aims to add to the findings in Study 1 supporting the use of implicit measurement in the detection and direction of bias towards attractive individuals in the context of employability. The predictions for the current study included: (1) the IRAP will indicate a bias towards attractive individuals as employable but not medium attractive individuals as unemployable; (2) that explicit measures will indicate a small but still present bias towards attractive individuals as more employable; and (3) participants will select the CVs of attractive individuals over the CVs of unattractive individuals. These predictions are based on both prior research conducted using explicit measures (Griffith & Langlois, 2006) as well as the results from Study 1.

Method

Participant details

Fifty-two individuals completed the study; three others participated but their data were discarded as they did not meet the IRAP inclusion criteria explained in the results section of study 1. The final study included 52 university students comprised of 18 males and 34 females ranging in age from 18-35, mean age 21.81. All participants completed an IRAP and four explicit questionnaires and rating scales. Participants were recruited from a participant pool comprised of first year undergraduate students who had indicated an interest in participating in research studies.

Ethical Considerations

Ethical issues for Study 2 were the same as Study 1, please see ethical considerations for Study 1 for full details.

Apparatus and Materials

The research was conducted in Maynooth University, in the psychology departments testing labs. The same testing rooms and laptop were used in study 2 as were used in study 1. The IRAP program controlled all stimulus presentation and recording of participant output (both participant time latency and accuracy). The stimuli presented by the IRAP was made up of six "good employee" words (e.g. responsible, intelligent) and six "bad employee" words (e.g. irresponsible, stupid). These words were retained from study 1. The IRAP also presented twelve pictures comprising of six male faces (three attractive and three unattractive) and six female faces (three attractive and three unattractive). The attractive images were the same as those used in study 1 while the unattractive images were changed for the purposes of the experiment. These unattractive images were chosen as more 'ambiguous' images than those in study 1 and had less extreme examples of unattractive images. In addition, the words "true" and "false" were presented as response options. The IRAP presented the stimuli in an identical manner to study 1 (see Table 1) with the only difference being the change in picture stimuli. The images used in study 2 can be found in Appendix 2. The same explicit measures that were used in study 1 were also presented in study 2 (the MIA, IJS and Likert scales; see Appendices 3 - 5).

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 examples 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.

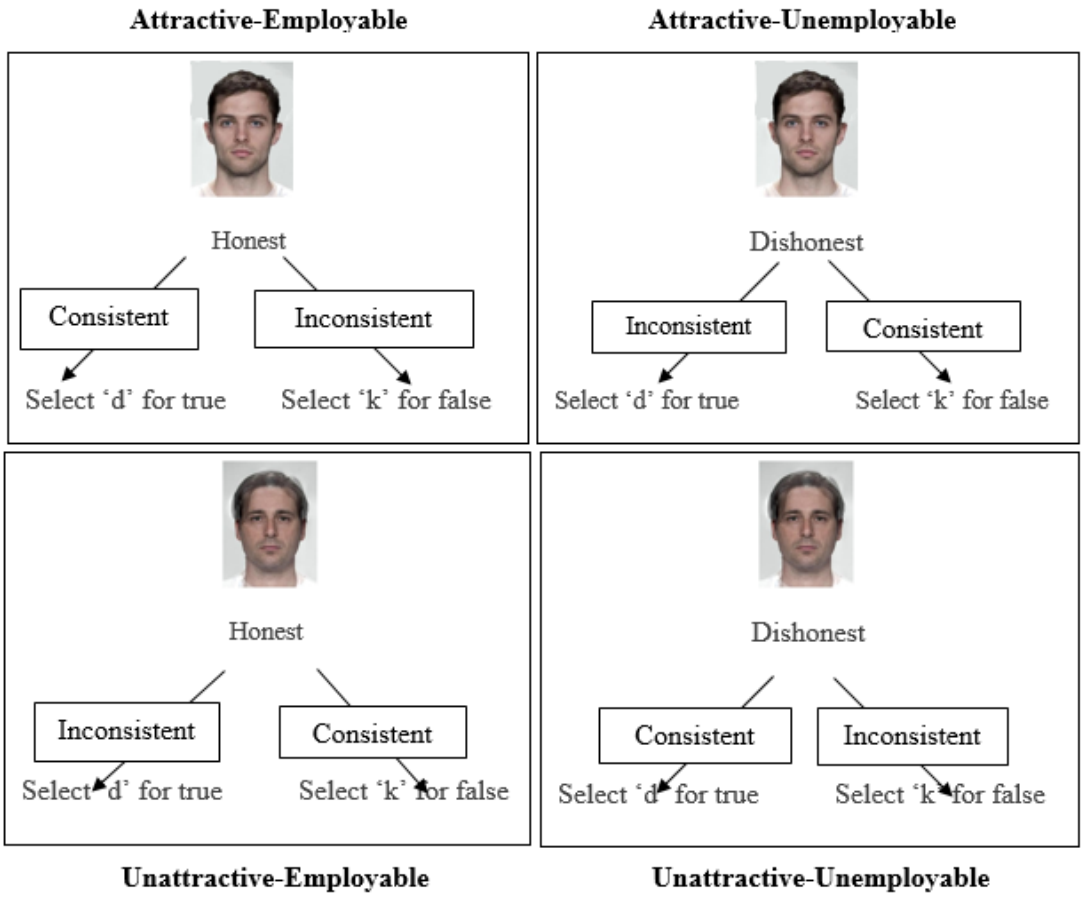


Figure 3. Representations of four IRAP trial-types. The superimposed arrows and labels indicate what would be considered a bias towards attractive-employable (consistent) or a bias towards unattractive-employable (inconsistent) response for each trial type. (These are for illustration purposes and did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

Results

Overview of data analysis

The data analysis for the explicit measures is presented first followed by the analysis of IRAP data. Statistical analysis includes comparisons of means for the attractive and unattractive explicit measures; IRAP data were subjected to, one sample t-tests, correlational analyses, analysis of variance tests (ANOVAs) and regression analysis. One sample t-tests were conducted to examine individual IRAP effects, correlational analysis were conducted to

examine relationships between the IRAP and explicit measures; an ANOVA was conducted to examine the effects of gender of participants on IRAP scores for each of the four trial types. Finally, a regression analysis was conducted to examine the predictive power of the IRAP in relation to a CV selection task (selecting between attractive vs. unattractive individuals).

Explicit measures

For the purpose of analysis and consistent with the scoring guide for each explicit measure: scores for the task attraction subscale of the Measures of Interpersonal Attraction (MIA) were calculated for each participant for both attractive and unattractive individuals (yielding two scores for this measure for each participant); similarly, scores for the Interpersonal Judgement Scale (IJS) were calculated for each participant for both attractive and unattractive individuals. Thus four scores for each participant were produced from the two scales (two for attractive and unattractive individuals). The means and standard deviations for each of the measures and conditions are presented in Table 4, along with the means and standard deviations of each of the four Likert scales. To examine if there was a difference in scores for each scale for attractive and unattractive faces, a number of paired sample t-tests and Wilcoxon signed rank tests (for normally and non-normally distributed scores respectively) were conducted. The results of these tests are summarised in Table 4. Predictably, attractive faces were rated as more attractive than the unattractive 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 unattractive faces, t (51) = -17.37, $p < .001$. Similarly, a Wilcoxon signed rank test revealed a statistically significant difference in scores for attractive and unattractive faces for the Employability Likert Scale, z = -5.83, $p < .001$. 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 = -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 unattractive faces, indicating that attractive faces were scored as more attractive and more employable than unattractive faces. Effect sizes ranged from .45 to .86 indicating a strong effect size using Cohen's (1988) criteria. These analyses therefore confirmed that the variables utilised in the study were functioning as planned.

Table 4: *Summary of Descriptive and Inferential Statistics for the Explicit Measures*

Explicit Measures	**Attractive Faces**		**Unattractive Faces**			
	M	*SD*	*M*	*SD*	*p score*	*Effect Size*
MIA	136.23	15.15	118.85	18.04	.000	.45
IJS	44.49	6.61	31.49	6.54	.000	.58
Attraction Likert Scale	13.75	3.31	-1.42	5.36	.000	.86
Employability Likert Scale	11.42	4.37	3.56	6.13	.000	.57

In addition to each of the explicit measures completed, a behaviour task was presented to each participant (CV selection). Although it was predicted that CV selections would be in line with the expected explicit results (i.e., that CVs with attractive pictures would be chosen over CVs with unattractive pictures) results were not consistent with this expectation. Data showed the participants selected more CVs with unattractive faces over attractive faces, with

33% of participants selected three attractive CVs faces (out of a possible four) and only 13% selected all attractive CVs. Out of the remaining 54% of participants, the majority of participants showed neutral responding, with 44% of participants selecting two attractive and two unattractive CVs, indicating no preference for attractive faces and 10% of participants selected three or more unattractive CVs over attractive faces. In summary, there was no consistent selection of attractive CVs over unattractive CVs.

IRAP data

The IRAP latency data for 52 participants were transformed into the four *D*-IRAP scores using the same transformation process employed in Study 1. As before with IRAP data a positive *D*-score was deemed pro-attractive and negative *D*-scores indicated pro-unattractive responding, however in order 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 & Barnes-Holmes, 2013). This inversion was to facilitate comparisons across trial types and does not alter the relative value of IRAP *D*-scores. All other presentation of IRAP *D*-scores are presented using uninverted scores for ease of understanding. On the graph (figure 4), scores above 0 indicate a pro-employable bias towards attractive images (i.e. 'pro-beauty' and 'anti-ugly') and scores below 0 indicate a pro-employable bias towards unattractive images (i.e. 'anti-beauty' and 'pro-ugly'). This data transformation yielded overall *D*-score that was positive, indicating that overall there was a bias favouring Rule A over Rule B (i.e. faster responding to affirm the relations attractive-employable/unattractive-unemployable). The *D* or "difference" scores reflect the difference in recorded mean latencies for both Rule A and Rule B; a positive *D*-score was deemed consistent with a pro-beauty bias and a negative *D*-score was deemed inconsistent with this stereotype. A single sample t-test revealed that the overall D-score (0.26) was statistically significantly different from 0, *t* (51)

= 7.56, p <.001, eta squared = 0.36. Thus the overall group data for the 52 participants demonstrated a pro-beauty bias that was statistically significant.

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 *Unattractive-Unemployable* (t (51) = 4.08, p < 0.001, eta = 0.14). The data for the remaining IRAP trial-type (*Unattractive-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 unattractive images as unemployable.

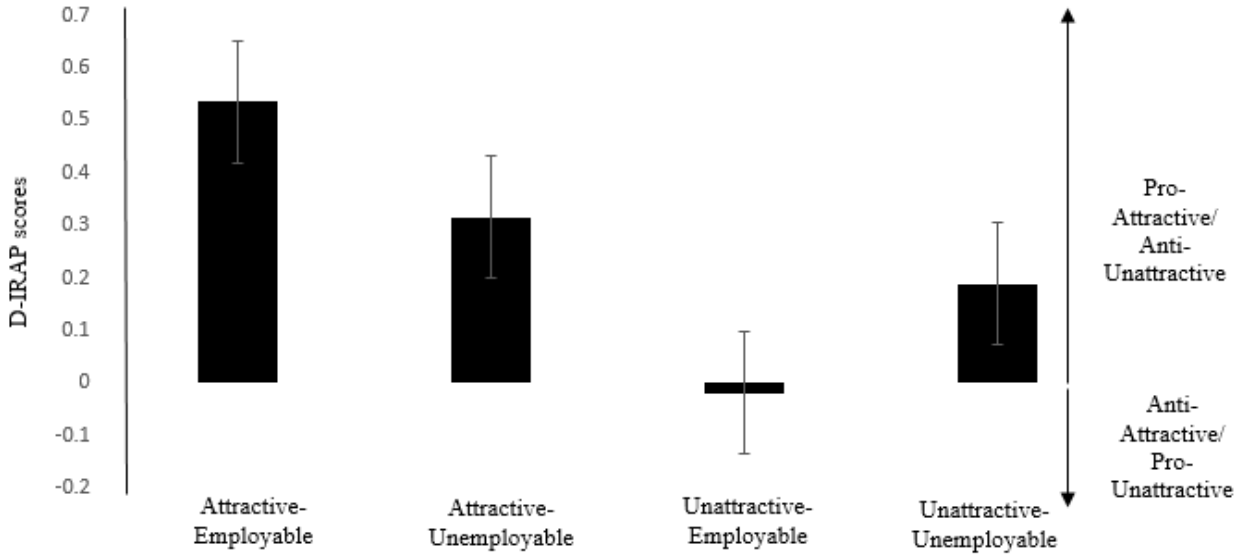


Figure 4: Mean *D*-IRAP scores for the 4 IRAP trial-types.

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 was combination of both pro-attractive and anti-attractive

Gender/ Trial-type 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 was 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 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).

Correlations between implicit and explicit measures

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 unattractive faces). Of the resulting correlations, none were statistically significant and these are presented in Table 5. In general, the scores on IRAP implicit measures did not predict scores on self-report measures.

Prediction of CV choice

In order to determine if the IRAP trial-type data predicted performance on a behavioural task (selecting from a CV choice between attractive/unattractive individuals), a multiple regression analysis was conducted. This involved using the overall mean *D*-IRAP

Table 5: *Correlations between IRAP trial-types and explicit measures.*

IRAP trial type	**Attractiveness Rating**		**Employability Rating**		**MIA**		**IJS**	
	Attractive	Unattractive	Attractive	Unattractive	Attractive	Unattractive	Attractive	Unattractive
Attractive-Employable	.11	-.12	.10	-.11	.01	-.12	.10	-.03
Attractive-Unemployable	-.00	-.00	.10	-.09	.24	-.15	-.23	-.02
Unattractive-Employable	.08	-.01	-.06	-.11	-.12	-.05	-.13	-.02
Unattractive-Unemployable	-.21	-.04	.06	-.02	.15	-.10	-.16	-.12

score calculated previously. 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 4 trial types *D*-IRAP scores were entered into the model. These also proved to be nonsignificant predictors of CV choice, with the model accounting for 2.7% of the variance, $p > .05$.

Summary

In summary, analysis of data in the explicit measures revealed that participants preferred attractive compared to unattractive facial images in the context of employability and this was the case for the IJS and the MIA measures, as well as on the Likert rating scale for employability; the difference in scores on all three explicit measures were statistically significant. For the behaviour task (select a CV with attractive/unattractive facial images), a small majority of participants choose CVs with unattractive images over CVs with attractive images. The Overall *D*-IRAP (N=52) score indicated a 'pro-beauty' bias; individual trial type analysis was showed a pro-attractive bias, there was a statistically significant bias indicating attractive images as employable and not-unemployable and a statistically significant bias indicating unattractive images as unemployable. Of the four trial types only the pro-unattractive-employable trial type did not reach statistical significance. This indicates the presence of a pro-beauty bias, and an anti-ugly bias in the current sample. Gender was not found to be influential. Correlation tests results were not statistically significant.

Discussion

The current study aimed to further the use of the IRAP as a measure of biases towards physical attractiveness and employability beyond its use with stimuli classed as high attractiveness or low attractiveness. The findings in study 2 indicate that using the IRAP with stimuli classed as medium attractiveness in comparison to high attractiveness individuals

provides a more nuanced understanding of the beauty bias effect. Overall, both explicit and implicit measures revealed that attractive faces were scored statistically significantly higher on both the attractiveness measures and employability measures. There was no consistent response pattern for CV choice in the behavioural task. The overall IRAP score indicated that responding was faster for rule A and confirmed the presence of a bias towards attractive individuals as employable and unattractive individuals as unemployable. Individual trial type analysis revealed a statistically significant bias towards attractive individuals as employable and not-unemployable, but also a statistically significant bias towards unattractive individuals as unemployable. This is a different result to Study 1, in that attractiveness was a benefit *and* unattractiveness was a disadvantage. This is an unexpected result, and a somewhat more negative result than those typically found in IRAP studies, but still gives a new interpretation of the beauty bias effect. As the two studies employed different participants, although similar, it is not possible to compare across groups. Therefore, it is not known if the different IRAP results found in studies 1 and 2 are due to IRAP differences or group differences.

Similar to study 1, an ANOVA revealed no effects for gender on IRAP scores which further supports the findings that sex of raters does not have influence on findings of employability and attractiveness. Although the study employed a larger sample size ($N_2 = 52$ as opposed to $N_1 = 24$) and as such should have more than sufficient power to employ a robust statistical measure such as the ANOVA (Vahey et al, 2015), a possible limitation is the unequal number of male and female participants. It also remains the case that the statistical power required for a 2x4 ANOVA may be too great for a sample of this size. This unequal number and the relatively small sample size may have confounded the ANOVA analysis and future research should endeavour to have equal numbers of male and female participants. No statistically significant correlations were found between the implicit and explicit measures although both had similar findings. This suggests that both explicit and implicit measures are

useful in measuring the beauty bias, however the IRAP may be useful in teasing out the full extent and directionality of the bias. Further research might be conducted into the lack of correlations, as both implicit and explicit measures indicated the same biases. Regression analysis revealed no statistically significant predictive power of the IRAP (both overall D and individual trial types) for CV choice. This is a limitation of the IRAP programme, as participants responded in a manner that indicated pro-beauty bias, however the IRAP failed to predict this. Further research might be conducted to examine if this is due to a failure of the IRAP programme or limited to this sample.

The current study confirmed and extended the results of study 1 – both the IRAP and explicit measures confirmed the presence of a bias towards attractive individuals as employable over unattractive individuals, even if the unattractive individuals were classified as medium attractiveness rather than low attractiveness. This is the first study to use implicit measures such as the IRAP to investigate the nuances of the beauty bias with different levels of attractiveness and findings such as those found in study 1 and the current study may have applied relevance in interviews and CV evaluations. Further research could look at direct comparisons between images of low, medium, and high attractiveness to investigate the true extent of the bias.

An important and often cited aspect of the beauty bias effect is the 'beauty is beastly' effect first investigated by Heilman and Saruwatari (1979). This may have contributed to the results found in both Study 1 and 2. Studies have found that attractiveness may be a disadvantage for attractive females when applying for high powered, stereotypical male positions but not for attractive males when applying for low powered, stereotypical female positions. This may have been contributing to the inconsistent findings on the IRAP, as female images may be seen as both employable and unemployable due to their gender and the

'beauty is beastly effect' (Cash, Gillen & Burns, 1974; Heilman & Saruwatari, 1979; Johnson et al, 2010; Paustian-Underdahl & Walker, 2016). Studies 1 and 2 did not separate images based on gender, only on the level of attractiveness. Study 3 separated the genders into male and female categories. The final study's aim was to investigate this finding using implicit measures, and the presence, if any, of a 'beauty is beastly' effect. This would be the first study to do so, and is an important extension of the literature in relation to beauty and hiring, and implicit measures such as the IRAP may allow a deeper and more complete understanding of the beauty is beastly phenomenon.

Chapter 4

Study 3

Using the IRAP to Measure Implicit 'Beauty is Beastly' Bias in Evaluations of Employability in Executive Roles for Male and Female Images

Studies 1 and 2 of the current research demonstrated that there was a consistent finding of participants rating attractive images as statistically significantly more employable than unattractive individuals of both low and medium attractiveness on explicit, self-report measures. This same bias was present during implicit testing using the IRAP, as both studies demonstrated a statistically significant pro-attractive bias. Study 3 aims to further investigate the beauty bias effect in the domain of employment by investigating the 'beauty is beasty' effect where attractive women are at disadvantage for managerial and stereotypical masculine jobs and when evaluated by the same sex (Heilman & Saruwatari, 1979; Cash, Gillen & Burns, 1977; Agthe et al, 2011). These studies found that female applicants were more rated as less likely to be hired or considered for managerial or 'male' sexed jobs when compared to male applicants. This finding has been reported in self-report measures, but has yet to be investigated using implicit measures such as the IRAP. As studies conducted using the IRAP to investigate beauty bias have comprised of mixed gender categories separated by attractiveness level and not gender (Murphy et al, 2015; Murphy et al, 2014), including the previous studies in the current research, no implicit measures have been used to evaluate the gender of images being rated. The effects of gender, where the paradox of female images being seen as both employable and unemployable is possible, may have influenced responding on both the IRAP and explicit measures and may provide some explanation for findings in the previous studies conducted in this research program. Similarly, to the beauty bias discussed in previous studies, using a IRAP to investigate this effect will allow for a more nuanced investigation and determining the directionality of the bias (e.g. is it pro-male, anti-female or some combination of both?).

The IRAP developed in the current study utilised both word and picture based stimuli. Six high powered executive job roles (CEO, manager, director, president, executive and CFO) and six low powered nonexecutive job roles (teacher, social worker, career, librarian,

secretary and assistant) appeared on screen, accompanied by one of either attractive male or attractive female pictures, and the response options 'true' or 'false'. These were presented in blocks of trials, which were deemed either *Rule A* (Pro-Male for executive job roles/Pro-Unattractive for nonexecutive job roles) or *Rule B* (Pro-Female for executive job roles/Pro-Male for nonexecutive job roles) – out of these rules, four trial types were involved: (i) *Male-Executive*, (ii) *Male-Nonexecutive*, (iii) *Female-Executive*, and (iv) *Female-Nonexecutive*.

Study 3 aims to build upon the previous studies by utilising the IRAP as a method for measuring and determining the direction of bias in the domain of physical attractiveness and employability. The predictions for the current study included: (1) that explicit measures will indicate a bias towards attractive male individuals for executive roles and female individuals for non-executive roles; (2) the IRAP will indicate a bias towards attractive male individuals for executive roles and female individuals for non-executive roles. These predictions are based on the findings of Heilman and Saruwatari (1979) and Cash, Gillen and Burns (1977).

Method

Participant details

Twenty-eight individuals completed the study; two others participated but their data was discarded as they did not meet the IRAP inclusion criteria explained in the results section of study 1. The final study included 28 university students comprised of 12 males and 16 females ranging in age from 18-35, mean age 22.68. All participants completed an IRAP and four explicit rating scales. Participants were recruited from a participant pool of first year undergraduate students who has indicated an interest in participating in research studies.

Ethical Considerations

Ethical issues for Study 3 were the same as Study 1, please see ethical considerations for Study 1 for full details.

Apparatus and Materials

The research was conducted onsite in Maynooth University, in the psychology department testing labs. The same testing rooms and laptop were used in the current study as in the previous studies. The IRAP program controlled all stimulus presentation and recording of participant output. The stimuli presented by the IRAP were made up of six "high-powered, executive" job titles (e.g. CEO, director) and six "low-powered, non-executive" job titles (e.g. teacher, secretary). These words were derived from a list of jobs provided by the United Kingdom's Office for National Statistics publication of women in the labour market. The IRAP also presented pictures comprising of attractive males and females as used in the previous studies (see Appendix 1 and 2). As in previous studies, the words "true" and "false" were presented as response options. The stimulus arrangements are presented in Table 6. As in the previous two studies, each participant's sex and age were recorded by the experimenter. A series of rating scales were presented to the participant, including a successfulness measure (Murphy et al, 2015) and two measures taken from Heilman and Saruwatari (1979) on the qualifications and recommendations of stimuli (see Appendices 6 and 7)

Table 6: *Stimulus arrangements presented by the IRAP*

Label 1	**Label 2**
CEO Manager Director President Executive CFO	Teacher Social Worker Carer Librarian Secretary Assistant
Sample deemed consistent with Label 1 Attractive male faces	**Sample deemed consistent with Label 2** Attractive female faces
Response Option 1 True	**Response Option 2** False

The successfulness measure consisted of a Likert scale consisting of the words '*successful'* and *'unsuccessful'* and participants scored pictures by circling a number from -3 (very unsuccessful) to +3 (very successful). The scales taken from Heilman and Saruwatari (1979) were two questions "how qualified do you think this applicant is for the position?" and "how likely would you recommend hiring this applicant?" accompanied with either an executive or non-executive position. Participants were asked to complete this scale by circling a number from -3 (very unqualified/not very likely) to +3 (very qualified and very likely).

Procedure

The procedure followed for the third study was identical to the procedure used in previous studies. The explicit measures were presented to the participant followed by the IRAP task. Examples of the four IRAP trial-types can be found in Figure 5. The only difference in the procedure of study 3 was that a CV choice behavioural task was not administered to participants.

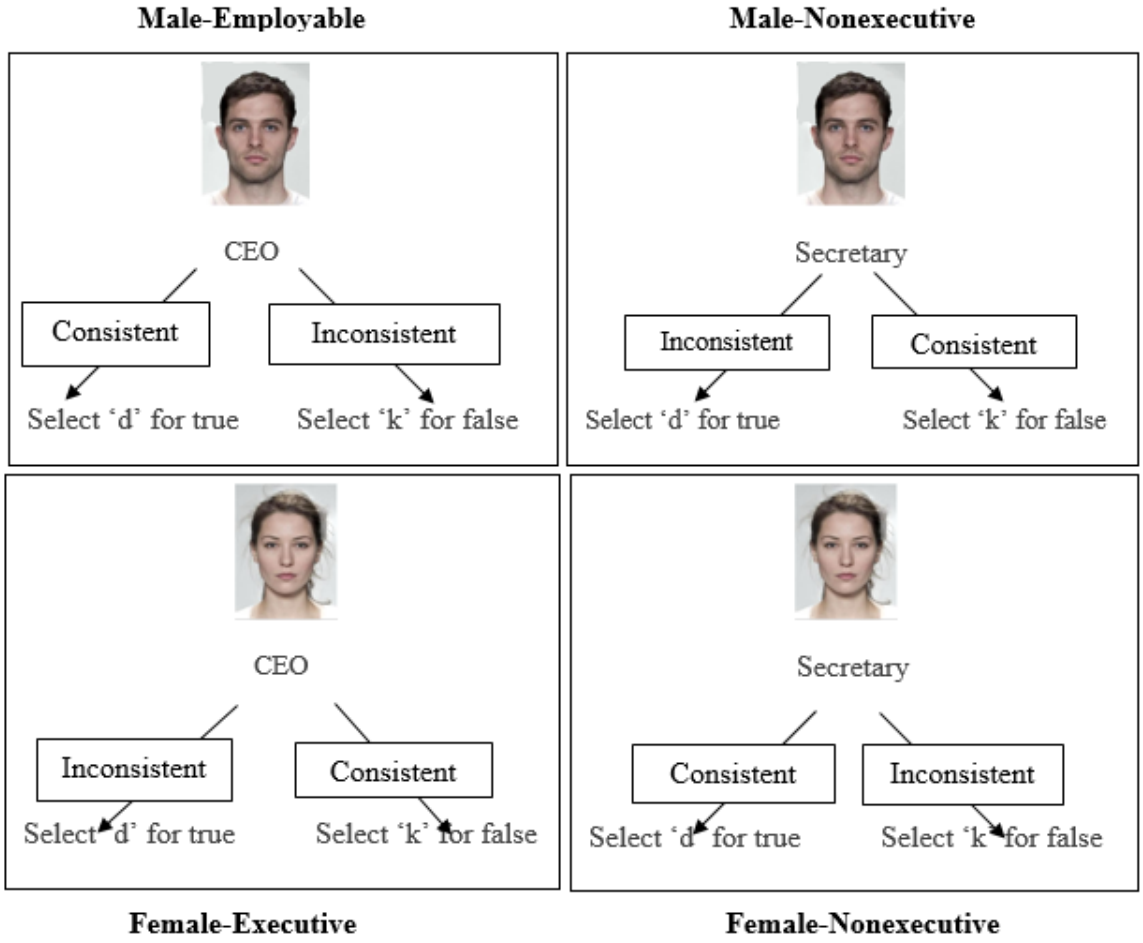


Figure 5. Representations of four IRAP trial-types. The superimposed arrows and labels indicate what would be considered a bias towards male-executive (consistent) or a bias towards female-executive (inconsistent) response for each trial type. (These are for illustration purposes only and did not appear on screen).

Results

Overview of data analysis

The data analysis for the explicit measures is presented first followed by the analysis of IRAP data, similarly to study 1 and 2. The analysis includes comparison of means for the male and female explicit measures, one sample t-tests, correlational analyses and an ANOVA. One sample t-tests were conducted to examine individual IRAP effects, correlational analysis were conducted to examine relationships between the IRAP and

explicit measures and an ANOVA was conducted to examine the effects of gender of participants on IRAP scores for each of the four trial types.

Explicit measures

For the purpose of analysis and consistent with the scoring guide for each explicit measure: scores for the successfulness measure, attractiveness measure and each of the Heilman and Saruwatari (1979) measure were calculated for each participant for both female and male individuals (yielding two scores for each measure for each participant). The means and standard deviations for each of the measures and conditions are presented in table 7. To examine if there was a difference in scores for each scale for male and female faces, a number of paired sample t-tests and Wilcoxon signed rank tests (for normally and non-normally distributed scores) were conducted. The results of these tests are summarised in table 7. Male faces were not rated as more attractive or successful than female faces, but were rated as more suited to executive jobs over nonexecutive jobs. For the Attraction Likert Scale, a Wilcoxon signed rank test indicated that there was no statistically significant difference in scores for male and female faces, $z = -1.61$, $p > .05$. For the Successfulness scale a paired sample t-test indicated there was no statistically significant difference in scores for male and female faces, $t\ (27) = .68$, $p > .05$. For executive jobs, t-tests indicated that there was a statistically significant difference for male and female faces, $t\ (27) = -3.19$, $p < .05$ and for nonexecutive jobs, t-tests indicated that there was a statistically significant difference for male and female faces, $t\ (27) = .47$, $p < .001$. All t-tests met the Bonferroni correction for *p*-values, $p < 0.012$. These tests indicated that there was no difference in scores for the successfulness or attractiveness measures for male and female faces, however scores were higher for males than females for executive job roles and scores were higher for females then males for nonexecutive job roles.

Table 7: *Summary of Descriptive and Inferential Statistics for the Explicit Measures*

Explicit Measures	**Male Faces**		**Female Faces**			
	M	*SD*	*M*	*SD*	*p score*	*Effect Size*
Successfulness	4.43	3.60	4.68	2.86	.500	.02
Attractiveness	5.43	4.23	6.75	1.80	.107	-.22
Executive Jobs	19.71	12.06	12.71	10.59	.004	.27
Nonexecutive Jobs	9.32	16.10	21.89	10.68	.000	.45

IRAP data

The IRAP data was transformed into the four *D*-IRAP scores using the same transformation process employed in previous studies. As before with IRAP data a positive *D*-score was deemed pro-male executive and negative *D*-scores indicated pro-female executive responding, however in order 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 & Barnes-Holmes, 2013). This inversion was to facilitate comparisons across trial types and does not alter the absolute value of IRAP *D*-scores. All other presentation of IRAP *D*-scores are presented using uninverted scores for ease of understanding. On the graph, scores above 0 indicate a bias towards male images in executive roles and scores below 0 indicate a bias towards female images in executive roles. This data transformation yields an overall *D*-score that was positive, indicating an overall bias towards Rule A over Rule B ("Males are

employed in executive roles and females are not"). Conducting a single sample t-test revealed that this score (0.27) was not statistically significantly different from 0, t (27) = .67, p > .05, eta squared = .01. A positive *D*-score indicate a pro-male executive bias and negative scores indicate a pro-male nonexecutive bias. Similarly, negative *D*-scores indicate a pro-female executive bias and positive scores indicate a pro-female nonexecutive bias.

The data from the 28 participants who completed the IRAP were included and the four mean *D*-IRAP scores for each trial type are presented in Figure 6 (see below). Only two of the trial types were statistically significantly different from 0: *Male-Executive* (t (27) = 4.67, p < .001, eta squared = 0.27); and *Female-Executive* (t (27) = -3.04, p < .05, eta squared = 0.15). The trial types *Male-Nonexecutive* and *Female-Nonexecutive* were not statistically significantly different from 0. These tests indicate that there were biases towards both males and females in executive roles, and none toward males and females in nonexecutive roles.

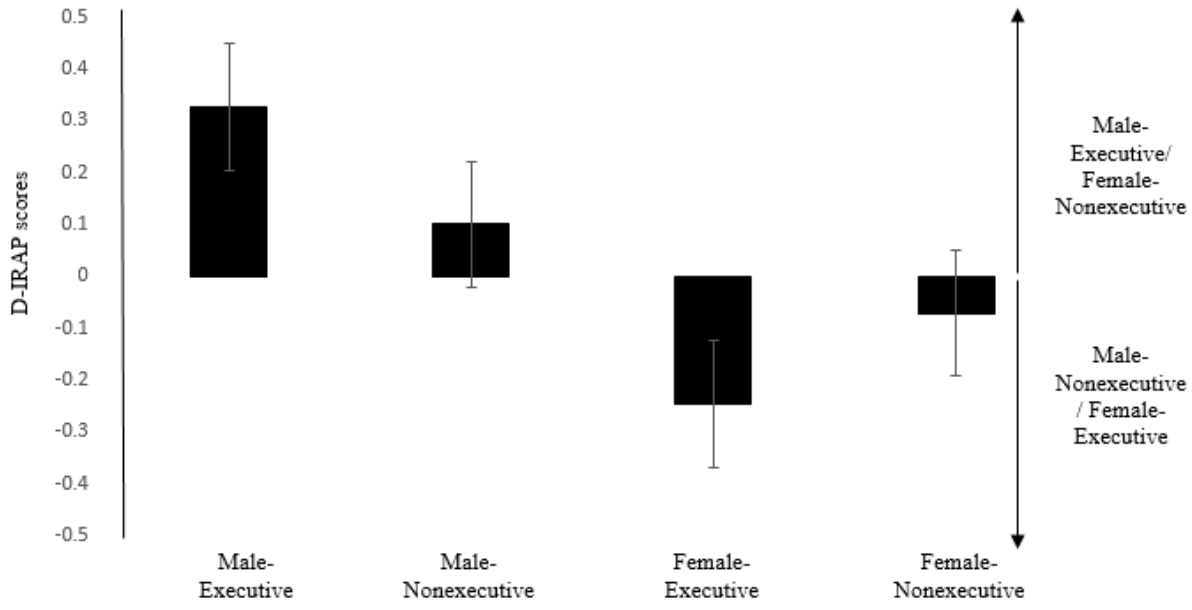


Figure 6: Mean *D*-IRAP scores for each trial type.

The 4 trial-type IRAP methodology allows for an examination of whether any bias shown is pro-male or anti-female or some combination of both; the current trial-type data show that the bias present is a combination of both pro-male and pro-female in executive jobs.

Gender/Trial type 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 = .88, F (3, 24) = 1.12, p = .36, partial eta squared = .12, and no statistically significant main effect for gender (p=0.99). This indicates that gender had no statistically significant effect on trial type scores and there was no overall differences between males and females. There was no statistically significant main effect for trial-type (Wilks Lambda = .75, F (3, 24) = 2.71, p = .07). Post hoc 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).

Correlations between implicit and explicit measures

Each of the *D*-IRAP scores for the four trial types were entered into a correlation matrix with the four explicit measures (see table 8). Of the resulting correlations between implicit and explicit measures, only one was statistically significant: the *Male-Nonexecutive* trial type correlated negatively with the attractiveness measure for male images, which indicates that an implicit bias towards males in a nonexecutive job role predicted that participants would rate male images as more unattractive. Interestingly, there was a number

Table 8: *Correlations between IRAP trial-types and explicit measures*

IRAP trial type	Successfulness		Attractiveness		Executive Jobs		Nonexecutive Jobs	
	Male	Female	Male	Female	Male	Female	Male	Female
Male-Executive	.05	.14	.04	.02	.09	.26	.00	.08
Male-Nonexecutive	.14	.24	-.47*	.00	.14	.03	-.34	.10
Female-Executive	.07	-.14	-.05	.32	.36	-.17	-.27	.00
Female-Nonexecutive	.23	.22	.03	.46	.22	.20	-.01	.18

*$p<0.05$

Table 9: *Correlations between explicit measures.*

	Successfulness		Attractiveness		Executive		Nonexecutive	
	Male	Female	Male	Female	Male	Female	Male	Female
Female successfulness	.84*		-.01	-.12	.29	.53*	.02	.32
Male successfulness		.84*	.20	.17	.50*	.55*	.11	.03*
Female attractiveness	.17	-.12	.47*		.48*	.28	.17	.10
Male attractiveness	.20	-.01		.47*	.28	.38*	.47*	.34
Female executive	.55*	.53*	.38*	.28	.48*		.68*	.71*
Male executive	.50*	.29	.28	.48*		.48*	.06	.68*
Female nonexecutive	.03*	.32	.34	.32	.68*	.71*	.51*	
Male nonexecutive	.11	.08	.47*	.17	.61	.68*		.51*

*$p<0.05$

of correlations between scales on the explicit measures (see table 9). For the female successfulness scale: as scores on this scale go up, so do scores on the male successfulness scale and the female executive scale. As scores on the male successfulness scale increase, so do scores on the male executive scale and the male nonexecutive scale. As scores on the female attractiveness scale go up, so too do scores on the male attractiveness scale and the male executive scale. Higher scores on the male attractiveness scale predict higher scores on the female executive scale and the male nonexecutive scale. As scores on the female executive scores increased, so too did the scores on the male executive, male nonexecutive, and female nonexecutive scores. As scores on the male executive scale increase so too do scores on the female nonexecutive. Finally, as scores on the female executive scores increase so too do scores on the male executive scales.

Summary

In summary, the explicit measures revealed that male and female faces did not score statistically significantly different on both the successfulness and attractiveness measures, indicating that participants did not rate either gender over the other. Participants did rate male faces statistically significantly higher for executive jobs than female faces and statistically significantly lower for nonexecutive jobs than female faces. This indicates that males were more likely to be associated with high-powered, executive roles. The Overall *D*-IRAP score indicated that there was a bias towards males in executive roles and females in non-executive roles, although this overall score was not statistically significant. Analysis of the four individual trial types indicated a statistically significant bias towards male and female images in executive roles (i.e. male and female images were both seen as executive), which contrasted with results on the explicit measures. There was no bias towards male or female faces in nonexecutive roles. Correlations between the implicit and explicit measures revealed only one statistically significant correlation between the male nonexecutive trial type and

ratings of attractiveness for males – implicit bias towards males in a nonexecutive job role predicted that participants would rate male images as more unattractive. An ANOVA conducted to examine the effects of gender on IRAP trial types indicated that there was no statistically significant effect.

Discussion

The current study aimed to investigate the presence of an implicit bias towards gender in executive roles, the 'beauty is beastly' effect and how this bias in comprised of (e.g. pro-male executive, anti-female executive, or some combination of the two) and the use of the IRAP as a measure of in this 'beauty is beastly' domain. Results from study 3 indicated that on explicit measures there was no difference between males and females on ratings of successfulness, but that for male images ratings for executive jobs were statistically significantly higher and ratings for nonexecutive jobs were statistically significantly lower than for female faces. This indicates that on explicit measures, attractive female images were at a disadvantage for executive jobs when compared to attractive male images. However, results from the IRAP indicate that there was a statistically significant pro-male executive bias *and* a statistically significant pro-female executive bias. The overall *D*-score was not statistically significant, indicating that participants did not respond in a pro-male executive/anti-female executive or pro-female executive/anti-male executive direction. There was no statistically significant bias towards non-executive jobs for either gender. This is seemingly incompatible with the scores on the explicit scales but the high number of positive correlations between the different explicit measures seems to suggest that participants were either responding in a socially desirable manner or in a rote manner throughout the questionnaire booklet due to the nature of the Likert scales. Again, as seen in the previous studies, an ANOVA revealed no statistically significant effect for gender on the IRAP trial type scores, which has demonstrated consistent support for the lack of sex differences on

scoring for employability, both male and female participants had broadly the same ratings across the IRAP trial types. However, the sample size in the current study may not have had enough statistical power for a 2x4 ANOVA to indicate a statistically significant gender effect.

IRAP results seem to indicate that there is strong bias positive towards male and female individuals in positions that are high-powered and executive, which is in direct contrast to the original Heilman and Saruwatari (1979) study as well as in partial contrast to more recent findings on the effect (Johnson et al, 2010; Paustian-Underdahl & Walker, 2016). It is important to note that the findings from recent studies appear to suggest that the beauty is beasty effect is strongest when attractive female applicants are applying for masculine-typed jobs where attractiveness is seen as unnecessary, because the current study used masculine-typed jobs where attractiveness is generally seen as helpful or necessary (e.g. CEO, manager) this may explain the lack of a beauty is beastly effect in the current study – further research using the IRAP to investigate this phenomenon should include masculine-typed jobs for which attractiveness is not seen as necessary (e.g. IT supervisor).

The data from study 3 were partially in line with predictions made in the introduction. Explicit measures did indicate that there was a bias towards attractive male individuals for executive roles and female individuals for non-executive roles. However, the IRAP indicated a statistically significant bias towards attractive male individuals for executive roles *and* attractive female individuals for executive roles. No statistically significant biases were found towards any gender and non-executive jobs. This is the first attempt to use the IRAP as an implicit measure to investigate the beauty is beastly effect, and attempted to contribute a new understanding of the directionality of the bias towards male and female attractiveness and employability that can be built upon and expanded in further studies. Implications of the current study and the previous studies will be discussed in detail in the following chapter.

Chapter 5

General Discussion

General Discussion

The current research aimed to investigate the beauty-bias effect in the context of employability. An implicit measure, the IRAP, was used to assess the presence and directionality of any biases towards physical attractiveness in the ratings of employability. Across three studies, a program of research was designed to investigate ratings of employability for stimuli of high and low attractiveness, medium attractiveness, and for male and female images in the context of executive and nonexecutive jobs. An IRAP was designed specifically for each study, and the implicit measure results were compared to commonly used explicit measures. Study 1 involved using an IRAP consisting of images of attractive and unattractive individuals and six 'good' employee words and six 'bad' employee words. Alongside the IRAP, two explicit measures were utilised the Measure of Interpersonal Attraction (MIA) and the Interpersonal Judgment Scale (IJS) and a brief behavioural task involving CV selection. Study 2 involved the same measures as Study 1, except that the stimuli used were images of high and medium attractiveness. Study 3, to investigate the presence of a 'beauty is beastly' effect, used an IRAP consisting of male and female attractive images and six 'executive' jobs and six 'nonexecutive' jobs. Alongside the IRAP explicit self-report Likert scales were taken from the original Heilman and Saruwatari (1979) study. 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. There was also a statistically significant bias towards attractive males and females in executive jobs compared to nonexecutive jobs. The final chapter will summarise the main findings of the three studies and discuss these findings in relation to the broader research field. Wider implications will be discussed, alongside limitations of the current research program, and suggestions for further research.

Summary of findings

Study 1 was the first study to utilise the IRAP as an implicit measure of physical attractiveness bias in the context of employability. Using undergraduate university students, an IRAP, four explicit measures (the Interpersonal Judgement Scale, the Measures of Interpersonal Attraction, employability and successful Likert Scales) and a behavioural task (a CV sorting task) were presented to twenty-four participants. Results from the IRAP indicated that there was a statistically significant bias towards attractive individuals as employable and not-unemployable, and there was no statistically significant bias towards unattractive individuals in either direction. On explicit measures, there were statistically significantly higher scores for attractive individuals than unattractive individuals across all measures. On the behavioural task there was no pattern to CV choice among participants, and multiple regression analysis indicated that there was no predictive power of IRAP *D*-scores (both overall and trial type) on the CV choice. Correlations between the explicit and implicit measures revealed that there were three statistically significant correlations: these correlations indicated that the implicit bias towards attractive individuals as not-unemployable predicted participants would rate attractive individuals as more attractive and also rate unattractive images higher on employability scales (IJS); correlations also indicated that an implicit bias towards unattractive individuals as unemployable indicated decreased scores on employability (MIA). Finally, an ANOVA revealed there was no influence of gender on IRAP trial type scores.

Overall, study 1 provided support for the use of the IRAP as a measurement of physically attractiveness bias towards employability and provided some much need directionality and nuance towards the phenomenon that explicit measurements were unable to ascertain. However, one possible criticism and a criticism of the wider physical attractive research, is that the stimuli used were of either high or low attractiveness (Langlois et al,

2000). As it is more likely for individuals to encounter people of medium, average attractiveness it may be that using stimuli such as this will affect scores on both the IRAP and explicit measures. The aim of the second study was to determine if the inclusion of medium attractive images would influence the effects found in Study 1. Study 2, therefore, replicated Study 1 in its procedure and measures except that images of high attractiveness and medium attractiveness individuals were employed instead of images of low attractiveness. The explicit measures and behavioural task were identical to that of study 1. A similar sample of participants were recruited from undergraduate university students, overall fifty-two individuals completed the study. Results from study 2 were similar to those of study 1, and confirmed the use of the IRAP as a viable measure of implicit biases in the domain of employability.

The results from study 2 again indicated a bias towards physically attractive individuals as employable on both the explicit and implicit measures, and was the first study to utilise the IRAP to investigate the effect that level of attractiveness play on the beauty-bias effect. Results from the IRAP indicate that there was a statistically significant bias towards attractive individuals as employable and not-unemployable, however in the current study there was also a statistically significant bias towards unattractive individuals as unemployable. This is a more negative result than that found in the previous study, as being attractive is an advantage *and* being unattractive is a disadvantage in employability measures. This is similar to results found by Griffin and Langlois (2006) where unattractiveness was consistent with a negativity bias. This is a result which is not possible to investigate in depth using explicit measures, as the full investigation into the directionality of the bias is not possible using self-report measures. This further underlines the use of the IRAP in the investigation of bias towards physical appearance in order to investigate the details of the bias. Similarly, the results from the explicit measures show that attractive images had

statistically significantly higher scores on the employability and attractiveness rating scales than unattractive faces. For the CV choice task a small majority of participants chose three or more attractive CVs over unattractive CVs (56%). Again, regression analysis revealed no statistically significant predictive power of the IRAP for CV choice. In study 2 the CVs were presented before the explicit measures so this may have allowed participants to respond without being cued as to the purposes of the experiment, however because the participants were different among the two studies it is not possible to directly compare across the studies. Correlations conducted between the implicit and explicit measures indicated no statistically significant correlations. Finally, an ANOVA conducted to determine if the gender of participants influenced scores on the IRAP indicated that there was no statistically significant gender effect.

The results of study 2 replicated those in study 1: both studies indicated a bias towards attractive individuals as employable and not-unemployable on the IRAP, as well as a bias towards attractive individuals on the explicit measures. Both studies indicated that there was no gender influence on the scores of the IRAP. The primary difference between the studies was the presence of the statistically significant bias of unattractive individuals as unemployable in the second study. Again, because the second study was comprised of a different sample it cannot be assumed that this result was not inherent to the second sample but nevertheless, it may be that the change in stimuli resulted in this change in results.

The aim of the final study was to investigate a further aspect of the beauty bias effect, the ‘beauty is beastly’ effect as coined by Heilman and Saruwatari (1979). Study 3 was designed to examine if bias exists towards female individuals in non-executive jobs and male individuals in executive jobs. An IRAP was designed using images of attractive and unattractive individuals and six executive and six nonexecutive jobs. Explicit measures used in the third study were taken from Heilman and Saruwatari (1979). Explicit measures

indicated that there was a bias towards males in executive roles and females in nonexecutive roles, with no statistically significant differences in ratings of successfulness or attractiveness. This is in line with results found previously (Johnson et al, 2010; Paustian-Underdahl & Walker, 2016; Cash, Gillen & Burn, 1977). However, results on the IRAP indicate that there were statistically significant biases towards both male and female images and executive jobs and surprisingly, there was no significant biases towards either gender for nonexecutive jobs. Correlational analysis revealed one statistically significant correlation between the explicit and implicit measures: an implicit bias towards males in nonexecutive roles was associated with ratings of male participants as unattractive. Surprisingly, there was several statistically significant positive correlations across the explicit measures. This suggests that as scores increased on one measure they increased across the other measures even if they seemed in direct opposition to each other. This seems to indicate that the explicit measures used may not be the best fit for measuring the exact biases found in participants, and the IRAP may provide a more useful tool for measuring these biases.

Specific Findings

The use of the IRAP as a measure of biases towards physical attractiveness and employability has not been well established in the wider literature of the physical attractiveness, with only two previous studies published utilising the IRAP as a measure of attractiveness bias (Murphy et al, 2014; Murphy et al, 2015). However, these studies only dealt with attractiveness in the domain of general successfulness and not specifically employability. The findings of the current research program display the possible use of the IRAP as a measure of employability in the context of physical attractiveness, and may have an advantage over the traditional self-report measures in investigating this effect.

The results of study 1 are in line with both the predictions of the research and the previous research done in this domain (Shahani-Denning, 2003; Watkins & Johnson, 2000; Cash & Kilcullen, 1985). These results support the 'what is beautiful is good' stereotype, with a bias towards attractive individuals being seen on both the implicit and explicit measures. However, the IRAP does allow a more nuanced investigation into this bias (Agthe et al, 2014). The ability to ascertain the directionality of the bias (that is, that it is pro-attractive and not anti-unattractive). The fact that the gender of participants had no effect on the scores obtained from the IRAP was also in line with similar research (Dipboye et al, 1977) and perhaps lend some credence to opponents of Feingold's (1990) argument that women might be more likely to conceal a bias towards attractive individuals as both male and female participants displayed the bias, and there was no statistically significant difference between the genders. Although there was no pattern to CV choice among participants, this may be due to the sequence in which they were presented to participants (after the explicit measures) which may have cued participants to the nature of the research. This may also be due to the content of the CVs, as although every effort was made to match the CV across experience and qualifications it may still be participants were unintentionally influenced by differences among the experience and qualifications.

Studies 1 and 2 demonstrated that the IRAP could consistently detect the presence of biases towards attractive individuals over unattractive individuals, and further could detect the specific direction that the bias lies in. The current studies found that across both studies there was a statistically significant bias towards attractive individuals as employable and not-unemployable. This indicates that it is an advantage to be attractive, and supports the 'what is beautiful is good' hypothesis (Dion, Berscheid & Walster, 1972; Eagly, et al, 1991; Cash, Gillen & Burns, 1977; Watkins & Johnston, 2000). However, the second study revealed a further aspect to the directionality of the bias in that there also was a statistically significant

bias towards unattractive individuals as unemployable. That is, that it is an advantage to be unattractive and a disadvantage to be unattractive. This is a finding that is unusual amongst the literature of physical attractiveness bias, as noted in Griffin and Langlois (2006) there exists a large bank of research underlying the 'beauty is good' bias but a much smaller research history of investigating if 'ugly is bad'. The results from study 2 indicate that there is, at least among the sample used, a bias towards 'ugly is bad' as well as 'beauty is good'. The IRAP may be a useful tool in further investigating the directionality of the bias as the nature of the four trial types can be manipulated and changed depending on the exact research question (Hussey et al, 2015) whereas self-report rating scales and questionnaires are static and largely unchanged across research questions. For example, as used in the current studies, one IRAP investigated biases towards individuals of high and low attractiveness and a second IRAP investigate biases towards individuals of high and medium attractiveness. Within these IRAPs it is possible to determine if biases are pro-attractive, anti-unattractive or some combination of the two.

The current research programme was the first to utilise both the IRAP and images of high and medium attractiveness in the domain of employability. As suggested by Langlois et al (2000), study 2 aimed to address a common criticism of physical attractiveness research by including images of medium attractiveness in order to investigate the effect fully. As suggested by both study 1 in the current research there again was a bias towards attractive individuals as both employable and not-unemployable, however in the current study unattractive individuals were also seen as unemployable. Although this result was not found in the first study, similar results were found in Griffin and Langlois (2006). Using explicit measures, they found a 'ugly is bad' bias and not the 'beauty is good' bias as so often found in the research, and this result was partially found in the current research. Although the explicit measures used in the study revealed broadly similar results to the first study, the

difference in the IRAP results from study 1 and 2 may be due to the difference in the level of attractiveness of the stimuli. Rubenstein, Langlois and Roggman (2002) suggested that attractive faces are perceived as more familiar than unattractive faces. By using faces of medium attractiveness in the study, the participants may have been responding using different relational frames than when they were required to judge faces of low attractiveness, and because the attractive images remained images of high attractiveness the same relational responses were utilised in each study (Barnes-Holmes et al, 2010). As the IRAP involved a shorter response time than the explicit measures, on the explicit measures the participants had more time to develop and extend the immediate relational response. This may explain why study 1 and 2 found different results on the IRAP but not on the explicit measures for unattractive images as unemployable. However, because the two studies involved different, albeit similar, samples of the student population it is not possible to directly compare across the groups, and any explanation regarding the differences in the results may be due to inherent differences within in the groups, although the differences in stimuli may be enough to explain the difference.

An interesting and unexpected result in study 3 of the current research was the bias towards both male and female images as executive. Study 3 aimed to investigate the presence of a 'beauty is beastly' effect where attractive women are seen as less employable for executive roles. The finding in the current research appears to be unusual amongst the 'beauty is beastly' research (Johnson et al, 2010; Paustian-Underdahl & Walker, 2016; Cash, Gillen & Burn, 1977). However, recent studies appear to suggest that the beauty is beasty effect is strongest when attractive female applicants are applying for masculine-typed jobs where attractiveness is seen as unnecessary (such as an IT manager). Study 3 used masculine-typed jobs where attractiveness is generally seen as helpful or necessary (e.g. CEO, manager), which may explain the lack of a beauty is beastly effect reported (Paustian-

Underdahl & Walker, 2016). Further, many of the studies investigating the 'beauty is beastly' effect are conducted in the United States and Canada, and it may be that the culture in Ireland may have some impact on the findings of the current study. Both Ireland and the EU put influence on the roles of women in the workplace (Men and Women in Ireland, 2013) and as undergraduate students, the participants may have come into more contact with these initiatives and literature. 44% of academic staff in Irish universities are female (HEA national review, 2016) and 80% of the healthcare and 75% of the education profession are female, with Ireland being the 9th country in the EU for gender equality and slightly above the average for the EU (Men and Women in Ireland, 2013). In an even wider context, Europe has several women in high-profile, high-power decisions, such as politicians like Theresa May and Angela Merkel who are both in the highest position of power in their respective countries. As this study used primarily a sample of university students it may be that they were particularly affected by these statistics, and as such these were automatic biases towards both males and females that were readily picked up by the immediacy of responding on the IRAP.

The IRAP results in study 3 were in direct opposition with findings on the explicit measures and the previous research done in the area. Paustian-Underdahl and Walker (2016) suggest a possible explanation for this: that when physical attractiveness is seen as helpful to the job role the 'beauty is beastly' effect is not as evident. The job types used in this study were jobs where attractiveness is seen as a bonus so this may explain this unpredicted result. The REC model (a RFT model for explaining the IRAP effect) may provide an explanation for the difference seen on the IRAP and the explicit measures (Barnes-Holmes et al, 2010). According to the model, IRAP scores derive from the immediate and brief relational responding of the participant where the most probable response emitted on an IRAP trial is determined by the verbal and nonverbal history of the participant. This response is often

determined by the history of the participant and the current contextual variables, and this immediate response is the one that will be emitted first most often under a time constraint. In consistent trials that match this immediate relation the response will be emitted quite quickly, in contrast inconsistent trials that do not match this immediate relation will require additional time to offset the automatic relation in order to respond accurately and so will be emitted less quickly. As the IRAP requires immediate relational responding, the latency between presenting the trial and the response will provide information on the immediate relation brought about by the participant and as a result, a bias that may be present (Barnes-Holmes et al, 2010). This also explains the differences in bias and bias strength between the IRAP and explicit measures: when completing explicit measures such as questionnaires there is no time limit that prevents additional relational responses that cohere with a wider relational network. For example, when asked to rate pictures of men and women as executive or nonexecutive the immediate relational response might be women = nonexecutive, but there would be additional relational networks such as "both people are similar", "both people look professional" or "discrimination based on gender is wrong". This relatively complex relational network is easier to engage with in the absence of a timed trial, whereas the IRAP requires an immediate response and so draws from the immediate relational response – or the automatic response as opposed to the carefully considered response (Barnes-Holmes et al, 2010). This finding may also be explained by the high number of positive correlations among the different explicit measures. As this indicates that as scores increase on one measure they increase across other, sometimes opposite measures it may be that participants were either responding in a socially desirable manner or in a rote manner throughout the questionnaire booklet due to the nature of the Likert scales. It is often possible with Likert scales that participants will simply circle the same number all the way down the sheet due to passive engagement with the questionnaire, unlike the IRAP which involves active engagement with

the program in order to successfully pass to the test blocks. However, it is important to note that this is only a possible explanation for the results found in the current study, and further research should be conducted before any definitive explanations are put forward.

Interestingly, across the three studies of the current research the correlations between the explicit and implicit measures were not statistically significant even though many of the explicit measures reported the same general attitudes. For study 1, there were three statistically significant correlations: these correlations indicated that the implicit bias towards attractive individuals as not-unemployable predicted participants would rate attractive individuals as more attractive and also rate unattractive images higher on employability scales; correlations also indicated that an implicit bias towards unattractive individuals as unemployable indicated decreased scores on employability. For study 2, there were no statistically significant correlations among the explicit and implicit measures. Finally, for study 3 there was only one statistically significant correlation between the explicit and implicit measures: an implicit bias towards males in nonexecutive roles was associated with ratings of male participants as unattractive. For studies 1 and 2 the explicit results were in line with the results found on the IRAP (that attractive individuals were statistically significantly more employable that unattractive individuals) but there was little to no correlations between the measures. This might indicate that the IRAP and the explicit measures are measuring similar but different aspects of the effect and may be explained using the REC model as outlined above. For study 3 results on the explicit measures and the IRAP diverged, and this would explain the lack of correlations between the two.

Wider implications of the current research

In the domain of beauty bias and employability, the current research developed an implicit measure for assessing the strength and direction of biases towards attractive and unattractive individuals. Results from the three studies indicate that there is a statistically

significant bias towards attractive individuals as employable, and some evidence to support that there exists a bias towards unattractive individuals as unemployable on both implicit and explicit measures. This suggests that in an Irish context, among undergraduate students who are likely to become management and may sit on interview panels, there is a strong evidence that the current participants are susceptible to the 'beauty is good' hypothesis. This has large implications for the employment sector, as it may be an important factor in the hiring decisions of interview panels etc. Ireland has legislation in place prohibiting other forms of discrimination, such as discriminating against race, gender, religion etc., but has no legislation or policies in place governing discrimination against unattractive individuals. Given that recent surveys have indicated that 60% of employers have searched for job candidates on their social network sites, which is a 49% increase over the ten years from 11% in 2006 (Harris Poll on behalf of Career Builder, 2016), and increased use of employment sites such as LinkedIn, it may be a possibility that the attractiveness bias may be affecting employment decisions before even an in-person interview, and even beyond in cases such as promotion (Talamas, Mavor & Perrett, 2016). The current study highlights the need for an accurate method of measuring biases such as this, and although the IRAP may be a measure more suited than explicit measure for assessing this, further research is needed replicating this and assessing the effects of intervention on these biases.

The current research had designed an IRAP specifically tailored to measure attractiveness and employability, in both the context of different level of attractiveness and gender difference. Most attractiveness bias research to date has primarily focused on explicit measures (Shahani-Denning, 2003; Johnson et al, 2010; Paustian-Underdahl & Walker, 2016) with very little research utilising implicit measures (Murphy et al, 2015; Murphy et al, 2014). Explicit measures may be more susceptible to self-report bias and socially desirable responding (Murphy et al, 2014) and due to the socially sensitive nature of biases it may be

difficult to fully investigate the directionality of the bias. As such, the IRAP may be a useful tool in order to measure automatic and immediate responses and may give researchers a more complete view of the beauty bias which may be resilient against the effects of social desirability. The IRAP also has a significant advantage over explicit measures as it can be used to ascertain the directionality of the bias towards physical attractiveness, that is it pro-attractive, anti-unattractive or a combination of the two. In the current research it was particularly useful to ascertain that participants had a bias towards attractive individuals as employable and not-unemployable by using the four trial-types that the IRAP consists of. The fact that the direction of the bias can be measured is an important advantage, as knowing the direction of the bias would allow the development of interventions and programmes designed to counteract the specific biases found. As the IRAP can rapidly capture the biases found in a specific population it may be a useful measure to aid in the designing of interventions, and can also detect the effects of an intervention rapidly (Cullen, Barnes-Holmes, Barnes-Holmes & Stewart, 2009).

An interesting implication of the current research is the lack of any gender bias in the current sample, both as target stimuli and as raters. All three studies in the current research indicated that there was no effect for participants' gender on scores on the IRAP, both for overall *D*-scores and trial type *D*-scores. This is supported by research such as Dipboye, Arvey and Terpstra (1977), who indicated that the gender of evaluators was not influential on scores, and both women and men are as likely to be as biased against unattractive individuals as each other. This was replicated in the current study, which showed similar findings. However, a confounding factor that may be influencing the lack of gender differences is the relatively low power of the sample size in order to conduct a 2x4 ANOVA. Although Vahey, Nicholson and Barnes-Holmes (2015) have found sufficient statistical power with limited sample sizes, the limited sample size in the current study may be masking the presence of any

effects that may be present. Similarly, study 3 of the current research indicated that there were no effects of the gender of stimuli for ratings of high powered, executive jobs. That is, that male and female stimuli were just as likely to be associated with executive jobs as each other. This has been discussed in more detail previously in the current chapter and some explanations have been put forward, however these findings have important implications for employability in Ireland. If there is no bias against female individuals in executive roles, and there is evidence that there could be biases against female individuals in traditionally male roles that are not associated with appearance, then interventions should be targeted towards roles such as those. Interestingly, there have been government initiatives aimed towards increasing the numbers of women entering science, technology, engineering and maths fields in the past number of years which indicates an awareness of the bias against women in such traditional 'male' roles.

Limitations of the current research and suggestions for future research

Due to the fact that this research was the first to investigate physical attractiveness in the domain of employability using the IRAP there are some limitations to the current study and suggestions for future research that necessitate further discussion.

Firstly, 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 generalisability of the current results. These were chosen due to the make-up of the participant pool, which was primarily composed of Caucasian individuals. However, it may be argued that the results obtained from the current study may not generalise 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. Further research in the

current domain should also consider investigating the effects that comparisons across high, medium and low attractive images might have on results of the IRAP. As seen in studies 1 and 2, there was consistent biases for attractive individuals as employable and in study 2 there was a statistically significant bias for unattractive individuals as unemployable. However, due to the nature of the participant sample making up the two studies being different, direct comparisons could not be made across the studies. Future research might compare directly across the three attractiveness levels in order to ascertain the full extent of the 'beauty is good' effect on employability.

Secondly, the current research recruited undergraduate university students as the participant sample in order to measure the presence of possible biases towards attractive individuals. University students were recruited for the current research program based on previous research that has indicated that university students' ratings of job applicants were nearly identical to those of professionals (Gordon, Slade & Schmitt, 1986), as well as 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.

Finally, an investigate into the IRAP as a predictor of actual behaviour was not investigated fully and inconclusive in the current research program. Although there was a brief behavioural task (CV sorting) employed in the current research which was largely neutral, and not indicative of a preference of one over the other, there was not a full investigation into the prediction of actual behaviour as employed in other, similar studies (such as packets including CVs, grades, references and more extensive hiring related questionnaires based on each CV; for example, Cash, Gillen & Burns, 1977). Further, logical

regression analyses revealed no predictive power for the IRAP on CV choice. Although other studies have reported predictive validity of the IRAP (Carpenter et al, 2012) the current study did not replicate these findings in the current domain. Although participants had indicated a pro-beauty bias on both the implicit and explicit measures, this bias was not evident in CV selection. The lack of predictive validity of the IRAP may indicate an issues between 'statistical significance' and 'meaningful significance', whereby statistically on a group level there exists the presence of a beauty-bias but this bias may not be sufficiently large or socially significant in order to affect 'real world' behaviour. More research is needed to ascertain the full extent of the predictive power of the IRAP in domains such as employability and the full extent of the significance of the IRAP effect. As the IRAP failed to predict CV choice in the studies, it may be that the IRAP effect is not of sufficient power to impact on 'real-world' behaviour and as such there may be more suitable measures of beauty-bias that should also be investigated.

In study 3, IRAP scores seem to indicate that there is strong bias towards male and female individuals in positions that are high-powered and executive, which is in direct contrast to the original Heilman and Saruwatari (1979) study. It is important to note that the findings from recent studies (Paustian-Underdahl & Walker, 2016) appear to suggest that the beauty is beasty effect is strongest when attractive female applicants are applying for masculine-typed jobs where attractiveness is seen as unnecessary, and because the current study used masculine-typed jobs where attractiveness is generally seen as helpful or necessary (e.g. CEO, manager) this may explain the lack of a beauty is beastly effect in the current study. Further research using the IRAP to investigate this phenomenon should include masculine-typed jobs for which attractiveness is not seen as necessary (e.g. IT supervisor) in order to investigate fully the 'beauty is beastly' effect for women in masculine roles.

Finally, as previously discussed, further research should also consider the effects of interventions on the scores of the IRAP and the biases associated with that. Previous research has indicated that brief counter-stereotypical information presented before the IRAP has reduced the bias previous shown (Cullen et al, 2009) and RFT indicates that additional relational training may be able to transform implicit biases. Future research could investigate the malleability of the IRAP scores based on these brief interventions, and the impact on actual behaviour.

Conclusion

The current research has developed and extended the use of an implicit measure, the IRAP, as a measurement of biases towards physical attractiveness and employability. The research demonstrated that the IRAP may have avoided the common disadvantages of explicit measures in order to measure the strength and directionality of biases towards attractiveness and employability. Results indicate that there is an attractiveness is employable bias, and tentative support for an unattractive is not employable bias. Consequently, the IRAP may have advantages for measure implicit biases over traditional, explicit measures. Further, the IRAP results revealed that there is also a significantly significant bias towards male and female images as executive and not nonexecutive in the domain of the 'beauty is beastly' effect and highlights the presence of attractiveness biases in Ireland. The current thesis extends both the existing attractiveness bias research and the existing IRAP research in using an implicit measure in measuring attractiveness bias in the domain of employability, and is the first to do so in the current field. Further research is still necessary in order to investigate the beauty bias effect fully, however the current research adds a dimension to the effect that has not been investigated and the IRAP may prove to be an important measure in the future of the attractiveness bias and employability.

References.

Agthe M, Spörrle M, & Maner JK (2011). Don't hate me because I'm beautiful: Anti -attractiveness bias in organizational evaluation and decision making. *Journal of Experimental Social Psychology, 46,* 1151-1154.

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., Smeets, P. M., Cullinan, V., & Leader, G. (2004). Relational frame theory and stimulus equivalence: Conceptual and procedural issues. *International Journal of Psychology and Psychological Therapy, 4,* 181-214

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.

Barnes-Holmes, D., Murphy, A., Barnes-Holmes, Y. & Stewart, I. (2010). The Implicit Relational Assessment Procedure: Exploring the impact of private versus public contexts and the response latency criterion on pro-white and anti-black stereotyping among white-Irish individuals. *The Psychological Record, 60,* 57-66.

Barnes-Holmes, D., Murtagh, L., Barnes-Holmes, Y., Stewart, I. (2010). Using the implicit association test and the implicit relational assessment procedure to measure attitudes toward meat and vegetables in vegetarians and meat-eaters. *The Psychological Record, 60,* 287-306.

Barnes, D., Lawlor, H., Smeets, P. M., & Roche, B. (1996). Stimulus equivalence and academic self-concept among mildly mentally handicapped and nonhandicapped children. *The Psychological Record, 46,* 87–107.

Beehr, T.A. & Gilmore, D.C. (1982). Applicant attractiveness as a perceived job-relevant variable in selection of management trainees. *Academy of Management Journal, 25,* 607-617.

Benson, P.L., Karabenick, S.A. & Lerner, R.M. (1976). Pretty pleases: The effects of physical attractiveness, race and sex on receiving help. *Journal of Experimental Social Psychology, 12*(5), 409-415.

Blair, I.V., Ma, J. & Lenton, A. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology, 81,* 828-841.

Bourisseau, W., Davis, O. L, & Yamamoto, K. (1965). Sense-impression responses to differing pictorial and verbal stimuli. *Educational Technology Research and Development, 13,* 249-25

Boyce, M.W. (1979) Physical attractiveness – a source of teacher bias? *Australian Journal of Teacher Education, 4*(1), 41-44.

Brigham, J.C. (1980). Limiting conditions of the "physical attractiveness stereotype": Attributes about divorce. *Journal of Research in Personality, 14,* 365-375.

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

Cash, T.F., Gillen, B. & Burns, D.S. (1977). Sexism and beautyism in personnel consultant decision making. *Journal of Applied Psychology, 62,* 301-310.

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.

Carpenter, K.M., Martinez, D., Vadhan, N.P., Barnes-Holmes, D., Nunes, E.V. (2012). Measures of attentional bias and relational responding are associated with behavioural treatment outcome for cocaine dependence. *American Journal of Drug and Alcohol Abuse, 38*(2), 146-154.

Castellow, W.A., Wuensch, K.L. & Moore, C.H. (1990). Effects of physical attractiveness of the plaintiff and defendant in sexual harassment judgements. *Journal of Social Behaviour and Personality, 5*, 547-562.

Central Statistics Office (2013). *Women and men in Ireland.* Dublin: Central Statistics Office Statistical Publication.

Chaiken, S. (1979). Communicative physical attractiveness and persuasion. *Journal of Personality and Social Psychology, 37*(8), 1387-1397.

Clifford, M.M. & Walster, E. (1973). The effect of physical attractiveness on teacher expectations. *Sociology of Education, 46*, 248-258.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cristofaro, TN, & Tamis-LeMonda, CS. (2008) Lessons in mother-child and father-child personal narratives in Latino families. In: McCabe A, Bailey A, Melzi G, editors. *Spanish-language narration and literacy: Culture, cognition, and emotion.* Cambridge: Cambridge University Press; pp. 54–91.

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.

Darby, B.W. & Jeffers, D. (1988) The effects of defendant and juror attractiveness on simulated courtroom trial decisions. *Social Behaviour and Personality, 16,* 39-50.

Dasgupta, N., McGhee, D. E., Greenwald, A. G., & Banaji, M. R. (2000). Automatic preference for White Americans: Eliminating the familiarity explanation. *Journal of Experimental Social Psychology, 36,* 316-328.

Dawson, D.L., Barnes-Holmes, D., Gresswell, D.M., Hart. A.J.P. & Gore, N.J. (2009). Assessing the implicit beliefs of sexual offenders using the IRAP: A first study. *Sexual Abuse: A Journal of Research and Treatment, 2,* 57-75.

De Houwer, J. (2002). The Implicit Association Test as a tool for studying dysfunctional associations in psychopathology: Strengths and limitations. *Journal of Behaviour Therapy and Experimental Psychiatry, 33,* 115-133.

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. (1972). Physical attractiveness and evaluation of children's transgressions. *Journal of Personal and Social Psychology, 24,* 207-213.

Dion, K.K. (1973). Young children's stereotyping of facial attractiveness. *Developmental Psychology, 9,* 183-188.

Dion, K.K. & Berscheid, E. (1971). *Physical attractiveness and sociometric choice in young children*. University of Minnesota.

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

Dion, K.K., Pak, A.W. & Dion, K.L. (1990). Stereotyping physical attractiveness: A sociocultural perspective. *Journal of Cross-Cultural Psychology, 21,* 378-398.

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.

Dixon, M. R., Rehfeldt, R. A., Zlomke, K. M., & Robinson, A. (2006). Exploring the development and dismantling of equivalence classes involving terrorist stimuli. *The Psychological Record, 56,* 83–103.

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.

Downs, A.C. & Harrison, S.K. (1985). Embarrassing age spots or just plain ugly? Physical attractiveness stereotyping as an instrument of sexism on American television commercials. *Sex Roles, 13,* 9-19.

Dunham, Y., Baron, A. S., & Banaji, M. R. (2006). From American city to Japanese village: A cross-cultural investigation of implicit race attitudes. *Child Development, 77,* 1268-1281.

Eagly, A.H., Ashmore, R.D., Makhijami, M.G. & Longo, L.C. (1991). What is beautiful is good, but…: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin, 110,* 109-128.

Efran, M.G. & Patterson, E.W.S. (1976). The politics of appearance. Unpublished manuscript, University of Toronto.

Feingold, A. (1990). Gender differences in effects of physical attractiveness on romantic attention: A comparison across five research paradigms. *Journal of Personality and Social Psychology, 59,* 981-993.

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.

Frieze, I.H., Olson, J.E. & Russel, J. (1991). Attractiveness and income for men and women in management. *Journal of Applied Social Psychology, 21,* 1039-1057.

Gordon, M.E., Slade, L.A., & Schmitt, N. (1986). The "science of the sophomore" revisited: From conjecture to empiricism. *Academy of Management Review, 11,* 191-207.

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

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D.S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self -concept. Psychological Review, *19,* 3-25.

Greenwald, A.G., McGhee, A.D. & Schwarz, J.L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and 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.

Gilmore, D.C., Beehr, T.A. & Love, K.G. (1986), Effects of applicant sex, applicant physical attractiveness, type of rater and type of job on interviewer decisions. *Journal of Occupational Psychology, 59,* 103-109.

Hamermesh, D.S. & Parkel, A. (1994). Beauty and the labour market. *American Economic Review, 84*, 1174-1194.

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

Hayes, S.C., Barnes-Holmes, D., Roche, B. (2001). *Relational Frame Theory: A Post -Skinnerian Account of Language and Cognition.* New York: Plenum.

HEA (2016). *National review of gender equality in Irish Higher Education Institutions.* Luxembourg: Publications Office of the European Union.

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 and Human Performance, 23*, 360-372.

Heilman, M.E. & Stopeck, M.H. (1985). Attractiveness and corporate success: Different causal attributes for males and females. *Journal of Applied Psychology, 70*, 379-388.

Hosoda, M., Coats, G., Stone-Romero, E. F., & Backus, C. A. (May, 1999). 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.

Jæger, M.M. (2011). "A thing of beauty is a joy forever"? Returns to physical attractiveness over the life course. *Social Forces, 89*(3), 983-1003.

Jackson, L.A. (1983). The influence of sex, physical attractiveness, sex role and occupational sex-linkage on perceptions of occupational suitability. *Journal of Applied Social Psychology, 13*(1), 31-44.

Jackson, L.A., Hunter, J.E. & Hodge, C.N. (1995). Physical attractiveness and intellectual competence: A meta-analytic review. *Social Psychology Quarterly, 58,* 108-122.

Jacobson, M.B. & Popovich, P.M. (1983). Victim attractiveness and perceptions of responsibility in an ambiguous rape case. *Psychology of Women Quarterly, 8*, 100 -104.

Johnson, S.K., Podratz, K.E., Dipboye, R.L. & Gibbons, E. (2010). Physical attractiveness in ratings of employment suitability: Tracking down the "Beauty is Beastly" effect. *Journal of Social Psychology, 150*(3), 301-318.

Kulka, R.A. & Kessler, J.B. (1978). Is justice really blind? The influence of litigant physical attractiveness on juridical judgement. *Journal of Applied Social Psychology, 8*(4), 366-381.

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.

Langlois, J.H., Ritter, J.M., Roggman, L.A. & Vaughn, L.S. (1991). Facial diversity and infant preferences for attractive faces. *Developmental Psychology, 27*, 79-84.

Lemm, K. & Banaji, M.R. (1999). Unconscious beliefs and attitudes about gender. In U. Pasero & F. Braun (Eds.), *Wahrnehmung und herstellung von geschlecht (Perceiving and performing gender*) (pp. 215-233). Opladen: Westdutscher Verlag.

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

McElroy, J.C. & DeCarol, T.E. (1999). Physical attractiveness on cognitive evaluations of saleswomen's performance. *Journal of Marketing Theory and Practice, 7*, 84-100.

McKenna, I.M., Barnes-Holmes, D., Barnes-Holmes, Y. & Stewart, I. (2007). Testing the fake-ability of the Implicit Relational Assessment Procedure (IRAP): The first study. *International Journal of Psychology and Psychological Therapy, 7*, 123-138.

Merwin, R. M. & Wilson, K. G. (2005). Preliminary findings on the effects of self-referring and evaluative stimuli on stimulus equivalence class formation. *The Psychological Record, 55*, 561–575.

Morrow, P.C. (1990). Physical attractiveness and selection decision making. *Journal of Management, 16*(1), 45-60.

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 and Psychological Therapy, 14*(3), 333-351.

Newman, J., Gabriel, S. & Jones, D.C. (2001). "There she goes now, Miss Ugly": A naturalistic study of appearance conversations among 5th and 6th grade students. *Paper presented at the biennial meeting of the society for research in child development, Minneapolis, Minnesota.*

Nicholson, E., & Barnes-Holmes, D. (2012). The Implicit Relational Assessment Procedure (IRAP) as a Measure of Spider Fear. *The Psychological Record, 62*, 263-278.

Nolan, J., Murphy, C., & Barnes-Holmes, D. (2013). Implicit relational assessment procedure and body-weight bias: Influence of gender on participants and targets. *The Psychological Record, 63*, 467-488.

Nosek, B. & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition*, *19*, 625 –666.

O'Reilly, A., Roche, B., Ruiz, M., Tyndall, I., & Gavin, A. (2012). The function acquisition speed test (FAST): A behaviour-analytic implicit test for assessing stimulus relations. *The Psychological Record, 62*, 507-528.

Office for National Statistics (2013). *Women in the labour market.* London: U.K. Office for National Statistics.

Paustian-Underdahl, S.C. & Walker, L.S. (2016). Revisiting the beauty is beastly effect: Examining when and why sex and attractiveness impact hiring judgments. *International Journal of Human Resource Management, 27*(10), 1034-1058.

Peris, T. S., Teachman, B. A., & Nosek, B. A. (2008). Implicit and explicit stigma of mental illness: Links to clinical care. *Journal of Nervous and Mental Disease, 196,* 752-760.

Ramsey, J.L. Langlois, J.H., Hoss, R.A., Rubenten, A.J. & Griffin, A.M. (2004). Origins of a stereotype: Categorization of facial attractiveness by 6-month-old infants. *Developmental Science, 7*, 201-211.

Rennels, J.L. & Langlois, J.H. (2014). Children's attractiveness, gender and race bias: A comparison of their strength and generality. *Child Development, 85*(4), 1401-1418.

Roddy, S., Stewart, I. & Barnes-Holmes, D. (2011). Facial reactions reveal that slim is good but fat is not bad: Implicit and explicit measures of body-size bias. *European Journal of Social Psychology, 41*, 688-694.

Rubenstein, A.J., Langlois, J.L. & Roggman, L. (2002). What makes a face attractive and why: The role of averageness in defining facial beauty. In G. Rhodes & L.A. Zebrowitz (Eds.), *Facial attractiveness: Evolutionary, cognitive, and social perspectives* (pp. 1-33). Ablex: Westport: CT.

Schneider, F.W., Gruman, J.A. & Coutts, L.M. (2012). *Applied Social Psychology: Understanding and Addressing Social and Practical Problems.* London: SAGE Publications.

Schmidt, K., & Nosek, B. A. (2010). Implicit (and explicit) racial attitudes barely changed during Barack Obama's presidential campaign and early presidency. *Journal of Experimental Social Psychology, 46,* 308-314.

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

Shahani-Denning, C., Dipboye, R. L., & Gehrlein, T. M. (1993). Attractiveness bias in the interview: Exploring the boundaries of an effect. *Basic and Applied Social Psychology*, *14*, 317-328.

Shahani-Denning, C. & Plumitallo, D. (1993). The influence of physical attractiveness and gender on disciplinary decisions. *Paper presented at the 5th annual convention of the American psychological society, Chicago, IL.*

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research*, *14*, 5-13.

Sigall, H. & Ostrove, N. (1975). Beautiful but dangerous: Effects of offender attractiveness and nature of the crime on juridical judgement. *Journal of Personality and Social Psychology, 31*(3), 410-414.

Smith, J. C., Bradley, M. M., & Lang, P. J. (2005). State anxiety and affective physiology: effects of sustained exposure to affective pictures. *Biological Psychology, 69* (3), 247-260

Spencer, B.A. & Taylor, G.S., (1988). Effects of facial attractiveness and gender on causal attributions of managerial performance. *Sex Roles, 19*(5), 273-285.

Stewart, J.E. (1980). Defendant's attractiveness as a factor in the outcome of trials. *Journal of Applied Social Psychology, 10*, 348-361.

Talamas, S.N., Mavor, K.I. & Perrett, D.I. (2016). Blinded by beauty: Attractiveness bias and accurate perception of academic performance. *PLoS ONE 11*(2), 1-18.

Teachman, B. A., & Brownell, K. (2001). Implicit anti-fat bias among health professionals: Is anyone immune? *International Journal of Obesity, 25*, 1525-1531.

Watkins, L.M. & Johnston, L. (2000). Screening job applicants: The impact of physical attractiveness and application quality. *International Journal of Selection and Assessment, 8*, 76-84.

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 and Psychological Therapy, 10*, 453-474.

Van man, E.S., Sultz, J.L., Nathan, L.R. & Warren, J. (2004). Racal discrimination by low -prejudiced whites. *Psychological Science, 15*, 711-714.

Vilathong, T.O., Lindner, N.M., & Nosek, B.A. (2010). Do unto others: Effects of priming the golden rule on buddhists' and christians' attitudes toward gay people. *Journal for the Scientific Study of Religion, 49*, 494-506.

Watt, A. W., Keenan, M., Barnes, D., & Cairns, E. (1991). Social categorization and stimulus equivalence. *The Psychological Record, 41*, 33–50.

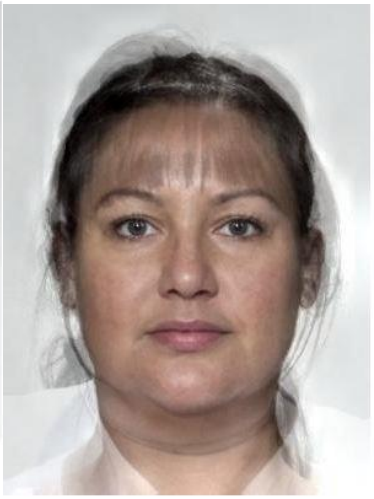
Appendix 1: IRAP picture stimuli study 1

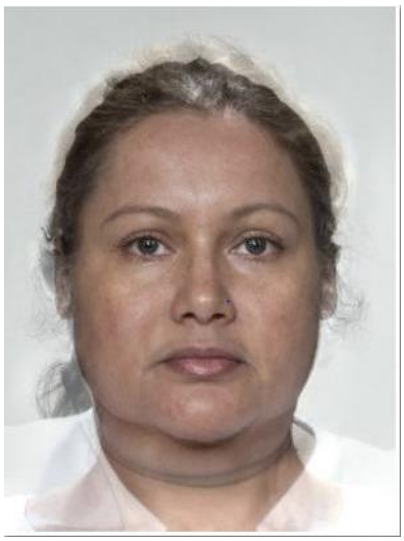
Attractive faces:

Unattractive faces:

Appendix 2: IRAP picture stimuli study 2

Medium Attractive Faces





Appendix 3: Interpersonal Judgement Scale

Please indicate the degree to which you agree with each statement by marking whether you:

Strongly disagree = 1; Disagree = 2; Neutral = 3; Agree = 4; Strongly Agree = 5.

1. To what degree is the target person:
    a. Intelligent?
    b. Moral?
    c. Knowledgeable of current events?
    d. Adjusted?
2. How much do you think you will like this person?
3. How much would you want to work with this person?

Appendix 4: Measure of Interpersonal Attraction

Please indicate the degree to which you agree with each statement by marking whether you:

Strongly disagree = 1; Disagree = 2; Neutral = 3; Agree = 4; Strongly Agree = 5.

1. He/she is a typical goof off when assigned a job to do. ____
2. You could count on him/her getting the job done. ____
3. I have confidence in his/her ability to get the job done. ____
4. If I wanted to get things done I could probably depend on him/her. ____
5. I couldn't get anything accomplished with him/her. ____
6. He/she would not be good to work with. ____

Appendix 5: Likert Scales

Please rate how attractive you find this image by circling one of the following options

Very Unattractive -3 -2 -1 0 1 2 3 Very Attractive

Please rate how employable you find this image by circling one of the following options

Very unemployable -3 -2 -1 0 1 2 3 Very employable

Appendix 6: Explicit measures (Heilman & Saruwatari, 1979)

All in all, how qualified do you think this applicant is for the position? **Manager**

Very unqualified -3 -2 -1 0 1 2 3 Very qualified

How likely would you be to recommend hiring this applicant?

Not very likely -3 -2 -1 0 1 2 3 Very likely

All in all, how qualified do you think this applicant is for the position? **Assistant**

Very unqualified -3 -2 -1 0 1 2 3 Very qualified

How likely would you be to recommend hiring this applicant?

Not very likely -3 -2 -1 0 1 2 3 Very likely

All in all, how qualified do you think this applicant is for the position? **Executive**

Very unqualified -3 -2 -1 0 1 2 3 Very qualified

How likely would you be to recommend hiring this applicant?

Not very likely -3 -2 -1 0 1 2 3 Very likely

All in all, how qualified do you think this applicant is for the position? **Secretary**

Very unqualified -3 -2 -1 0 1 2 3 Very qualified

How likely would you be to recommend hiring this applicant?

Not very likely -3 -2 -1 0 1 2 3 Very likely

Appendix 7: Likert scales

Please rate how attractive you find this person by circling one of the following options

Very Unattractive	-3	-2	-1	0	1	2	3	Very Attractive

Please rate how successful you think this person would be by circling one of the following options

Very unsuccessful	-3	-2	-1	0	1	2	3	Very successful

Appendix 8: Information Sheet

Details about the researcher:

Rachel Murphy, B.A. Psych, doctoral student

Contact number: REMOVED

Email: rachel.murphy.2012@nuim.ie

Details about supervisor:

Dr. Michelle Kelly

Contact number: 01 474 7470

Email: michelle.e.kelly@nuim.ie

What is the purpose of the research?

Attractiveness bias is thought to be an important factor in a number of different social contexts and have an impact on interpersonal relationships (such as favourable first impressions, ratings on performances of singers or musicians, how persuasive a person is, how intelligent they are thought to be, and even what sentence they receive in a court). The current research aims to answer the following questions: is there is a difference in attraction bias in implicit and explicit measures? In what direction is this bias found? And can this bias be changed or reduced?

Explicit measures are measures that ask the participant directly about the subject, such as questionnaires, surveys and interviews. Explicit measures allow the participant to think about their answer for as long as they wish, and are measures that you deliberately think about and report.
Implicit measures are measures designed to interpret a participant's automatic thoughts about a subject, without the participant consciously deciding what to respond. These measures usually consist of computer tasks, where the computer program measures all the information required.

What will the research involve?

The research will involve two session with the experimenter, one session with the experimenter now and another session in approximately six weeks. Firstly, the researcher will administer a number of pen and paper questionnaires that you will be required to complete. After this you will be given a task to sort CVs and pick a candidate to fill a job vacancy. These questionnaires should not take longer than twenty minutes. After these questionnaires are completed you will be asked to complete a computerised task called the IRAP. You will be given full instructions on how to complete this program beforehand, and any help you might need to complete the program. This computer program can be difficult, but you will be given all the information and help you need. This section of the research should take no more than forty minutes.
The second session will be similar to the first session, where you will be given some information to read before again completely the IRAP. Similarly, you will be given all the information and help needed to complete the program and it should take no longer than forty minutes.

When will the research be conducted?

Research will commence in September 2015 and will end no later than January 2017. Participation can occur at any time throughout September 2015 – January 2016 or March 2016 – September 2016 depending on a time that suits both you and the researcher. The research session should take no more than an hour, depending on how quickly the tasks are completed.

Where will the research be conducted?

The research will be carried out at Maynooth University, Maynooth, Co. Kildare.

What if I don't want to participate?

If you prefer not to take part in this research, please be assured that there is no obligation to participate nor is there any consequences for not participating. If you do volunteer to participate please be assured that you will be asked if you wish to continue halfway through your experiment session. If you consent now and later change your mind, due to the fact that all data will be anonymous and assigned a code name that cannot be matched to you after your experiment session has finished it will not be possible to withdraw your data after the experiment session has finished. You can remove your data at any time up until the point that the experiment session has finished by not continuing the experiment and contacting the researcher immediately.

How will my data be kept safe?

Your data will be assigned a code number and will be stored in a protected hard drive kept by the researcher for ten years and only accessable by the researcher. No key to connect participants' real names with their code numbers will be kept, by the researcher or any other individual. After a period of ten years the data will be destroyed – the hard drive will be formatted and overwritten. Individual results from the IRAP program or the pen and paper questionnaires will not be made available to participants, as the researcher cannot associate any participant with their data. Your data will be included in a formal write up of the experiment for the researcher's doctoral thesis, and may be used in future publications or presentations. Your data will not be used for any other purposes.

Thank you for taking the time you read this information sheet, if you have any further questions or need more information on any part of this sheet please do not hesitate to ask.

Appendix 9: Consent Form

INFORMED CONSENT FORM FOR PARTICIPATION IN DOCTORAL RESEARCH

Title: Exploring attractiveness bias using the Implicit Relational Assessment Procedure (IRAP).

The current research will be conducted by Rachel Murphy, B.A. (Hons) Psych., who is a doctoral student at the Department of Psychology, National University of Ireland, Maynooth, Co. Kildare, and can be contacted via telephone at 086 732 0341 or by email at rachel.murphy.2012@nuim.ie. The research will be supervised by Dr. Michelle Kelly who is a lecturer and researcher at Maynooth University and can be contacted at Michelle.E.Kelly@nuim.ie.

In agreeing to participate in the research study being carried out, I _____, understand the following:

- In conducting the current research project both the research student and the research supervisor are responsible for following the ethical guidelines given out by the Psychological Society of Ireland and the Behaviour Analyst Certification Board in all dealings with me.
- That the information sheet attached will tell me what procedures will be completed with me.
- My identity will not be provided in any subsequent publication of data, or any further presentation. All data will be assigned code numbers and no key identifying each participant's real name with their code number will be kept. All data will be stored on a protected hard drive kept by the researcher for a period of ten years after which it will be deleted, with the hard drive being formatted and overwritten.
- Due to the fact that the researcher will not be able to identify any individual participant's data, I will not be able to withdraw my data from the study after the experiment session is completed. Up until that point, I will be reminded that I can refuse to participate or withdraw from the study at any stage without any negative consequences.
- The study will include both pen and paper questionnaires and a computerised task know as the IRAP with the aim of investigating attractiveness bias and views on employability .
- **Important:** If I have a history of any seizure disorder or has experienced negative effects when viewing a computer or television screen, I should notify the researcher of this fact and I should carefully consider my participation in this research study.

I confirm that I have read and understand the consent form and the accompanying information sheet and that I agree to participate in this study.

Signed:

_____ Participant

_____ Researcher

_____ Date

If during your participation in this study you feel the information and guidelines that you were given have been neglected or disregarded in any way, or if you are unhappy about the process, please contact the Chairman of the Research Ethics Subcommittee, Dr Andrew Coogan. Tel: (01) 7086624 Email: andrew.coogan@nuim.ie. Please be assured that your concerns will be dealt with in a sensitive manner.