An Investigation into the Relationship between the Gender Binary and Occupational Discrimination Using the Implicit Relational Assessment Procedure

Aoife Cartwright[1], Ian Hussey[2], Bryan Roche[1], Jenny Dunne[1] & Carol Murphy[1]

[1] Maynooth University, Ireland, [2] Ghent University, Belgium

Reference:

Cartwright, Hussey, Roche, Dunne & Murphy (2017). An Investigation into the Relationship between the Gender Binary and Occupational Discrimination Using the Implicit Relational Assessment Procedure. *The Psychological Record, 67*(1) doi: 10.1007/s40732-016-0212-1

Author note:

Correspondence can be addressed to Aoife Cartwright, aoife.m.cartwright@gmail.com

Abstract

The social construction of gender-as-binary plays an important role within many contemporary theories of gender inequality. However, to date, the field of psychology has struggled with the operationalization and assessment of binarist ideologies. The current article proposes a technical framework for the analysis of the gender binary and assesses the suitability of the Implicit Relational Assessment Procedure (IRAP) as a measure of binarist gender beliefs. Forty-seven undergraduate students (26 female; M_{age} = 23.84) completed two IRAPs assessing the coordination of certain traits exclusively with women and others exclusively with men. Effects found on the IRAP were in the expected direction (i.e., relating men but not women with certain traits and women but not men with other traits). In addition, the traits ascribed to men within the IRAP were evaluated as more hirable by a large majority of participants (83%) on an explicit preference task. The results therefore support the arguments that, first, gender traits do seem to be framed oppositionally in language and, second, this binary may underpin existing gender hierarchies in certain contexts.

Keywords: gender binarism, Relational Frame Theory, gender discrimination, Implicit Relational Assessment Procedure

An investigation into the relationship between the gender binary and occupational discrimination using the Implicit Relational Assessment Procedure

Feminist theorists have long argued that the social construction of gender in terms of binaries is problematic (e.g., Butler & Butler, 2002). Dichotomizing gender into two discrete categories not only imposes strictures on male and female behavior (i.e., in terms of masculine and feminine "gender roles"), but it has been argued that the prioritization of men within androcentric or patriarchal systems is in part predicated on an ideology of real and complete gender difference (Bem, 1993; Kimmel, 2000). Put simply, if men and women are framed as fundamental opposites (i.e., what is "male" is also that which is "not-female" and vice versa; Connell & Messerschmidt, 2005), and men are more readily coordinated with certain socially-valued traits (e.g., leadership, see Eagly & Carli, 2007), then it is not only that men are leaders, but that women, by definition, are not.

To date, most of these arguments have been derived from mainstream feminist (e.g., the concept of gender performativity: Butler, 1990) and social-cognitive perspectives (e.g., the role congruity hypothesis: Eagly & Karau, 2002; the backlash effect: Rudman & Glick, 2001). Though neither field can be considered monolithic in its approach to gender relations, these accounts typically model inequality in terms of cognitive or cultural belief systems, attitudes, or ideologies. However, while these accounts are undoubtedly useful for researching and making visible the different ways in which society treats men and women, models anchored around such intrinsically mentalistic or sociological concepts do not readily lend themselves to a technical empirical analysis (see O'Reilly, Roche, & Cartwright, 2014).

The functional account of verbal behavior outlined by Relational Frame Theory (RFT: Hayes, Barnes-Holmes, & Roche, 2001) might constitute a viable behavioral alternative to the accounts mentioned above. RFT conceives of language in terms of networks containing

multiple and potentially infinitely complex relations between stimuli (e.g., same, opposite, more than, etc.; see Roche & Dymond, 2013 for a review of the evidence). From its earliest days, RFT conceived of an "attitude" as a network of trained and derived stimulus relations, established within an individual's verbal history (Grey & Barnes, 1996; Moxon, Keenan, & Hine, 1993; Roche, Barnes-Holmes, Barnes-Holmes, Stewart, & O'Hora, 2002; Watt, Keenan, Barnes, & Cairns, 1991). By modeling language and social processes in this way, RFT may provide a conceptual framework that allows us to integrate feminist arguments and social-cognitive conceptualizations into a technical analysis of social processes (i.e., in terms of networks of stimulus relations).

An RFT perspective of binarist gender ideologies might proceed as follows: the classes *women* and *men* are diametrically related; that is, they participate in a frame of opposition and/or distinction. In addition, *women* and *men* are coordinated with distinct sets of roles, interests, and abilities (i.e., *men* is the same as *masculine* and *women* is the same as *feminine*). The social practice of verbally "constructing" gender in this way can be readily observed in many Western cultures. For instance, young children are often explicitly trained (via parental instruction, children's reading materials, etc.) that "boy" and "girl" represent distinct, oppositional categories (see Gelman, 2005 for a comprehensive review). In other words, children learn from infancy to frame gender categories oppositionally, a relation that is arguably reified in increasingly complex ways over the lifespan (e.g., through society's continued discomfort with or intolerance of gender ambiguity; see Kimmel, 2007). Additionally, through participation with the dominant verbal community and interaction with popular culture, children observe the explicit coordination of these categories with a diverse range of attributes, behaviors, and topographical features, which effectively synchronizes biological "sex" with a range of culturally constructed gender roles (e.g., boys have short hair

and like cars, girls have long hair and like dolls; Kessler, Ashenden, Connell, & Dowsett, 1985; Witt, 1997).

In a similar vein, it could be argued that although individuals in Western cultures are not necessarily taught that women and men are inherently ill-suited for certain roles, they are routinely exposed to an array of verbal contingencies that support gender inequality in numerous forms. For instance, as noted previously, one possible consequence of framing men and women as relational opposites with distinct attributes is that roles related to one gender thus become oppositional to the other (i.e., opposition relations are derived and functions are transformed in accordance with the network). For example, if a young boy perceives his sister taking great interest in personal grooming, the oppositional gender binary relation may transform the function of personal grooming such that he perceives it as inappropriate for him to do the same as a boy. Similarly, an individual may or may not be explicitly taught that femininity is mutually exclusive with certain socially valued traits (e.g., success in a business context), but the gender network maintained by the verbal community implicitly supports the coordination between maleness and business competency and, more importantly, femaleness and a lack of competency in this area (see Figure 1).

[Figure 1 about here]

Modeling belief systems or attitudes in these relational terms allows for a technical analysis of the relationship between culture (i.e., the verbal community), language, and systemic gender inequality. Moreover, given that an RFT approach specifies both a behavioral process (relational responding) and a verbal network of interest (male-female

opposition relations and men-masculine/women-feminine coordination relations), this framework easily lends itself to an empirical investigation using procedures designed for assessing taught and derived stimulus relations. While basic stimulus matching procedures (e.g., matching-to-sample) have been some use in this regard (Moxon et al., 1993; e.g., Watt et al., 1991), more subtle and potentially sensitive procedures have been developed precisely for assessing stimulus relations that are socially sensitive. These include the widely known Implicit Association Test (IAT: Greenwald, McGhee, & Schwartz, 1998), the Function Acquisition Speed Test (FAST: O'Reilly, Roche, Ruiz, Tyndall, & Gavin, 2012), and the Implicit Relational Assessment Procedure (IRAP: Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010). Though a full review of these measures is beyond the scope of this paper, in brief, these "conflicting relations" paradigms typically compare response patterns to different stimulus relation pairings (e.g., the stimulus classes *men* and *women* with stereotypically masculine and feminine traits) across two different types of test blocks (e.g., men-masculine/women-feminine and women-feminine/men-masculine). Differences in response latency, accuracy, or fluency are then compared across the two blocks, with the differential assumed to reflect levels of consistency with that individual's verbal history (see De Houwer & Moors, 2010, for a more detailed description of these test formats; and Hussey, Barnes-Holmes, & Barnes-Holmes, 2015 for a discussion of the utility of such measures within research on derived relations).

Measures such as the FAST and the IAT have already been used to assess histories of relating men and women with distinct sets of traits or attributes (e.g., Cartwright, Roche, Gogarty, O'Reilly, & Stewart, 2016; Rudman & Glick, 2001). However, the IRAP does have one pragmatic benefit in this context in that it allows for the separation of latency differentials at the trial-type level. Specifically, given that each trial within the IRAP contains a single relation (e.g., the words "men" and "nurturing" as well as response options such as

true or *false*), IRAP scores can be calculated to assess response-time differentials for the particular trial type across the two types of test blocks (e.g., speed of responding to men-feminine-*true* relative to men-feminine-*false*). As such, the IRAP can produce four distinct latency differential scores, in this case: women-feminine, women-masculine, men-feminine, men-masculine. Unlike the FAST and the IAT, therefore, the IRAP can identify not only that certain traits are perceived as male but also that they are perceived as *not-female*, a distinction which may be conceptually important. A small body of research has already attested to the measure's suitability in this domain, by using the IRAP to measure, for example, relations between men/women and gendered household chores (i.e., chopping wood vs. cooking; Drake, Kellum, Wilson, Luoma, Weinstein & Adams, 2010) and between boys/girls and gendered toys (i.e., dolls vs toy cars; Rabelo, Bortoloti, & Souza, 2014).

This study will utilize the IRAP to assess binarist men-masculine and women-feminine verbal relations; that is, the ascription of certain attributes to women and not men and others to men and not women. To control for and assess any effects based on stimulus valence, two separate IRAPs will be employed: one for positively valenced traits and another for negatively valenced traits. To explore whether the traits related to the stimulus class "man/male" are indeed more socially valuable than those related to "woman/female" (i.e., that there is a hierarchy within the binary), this study will also employ a short hypothetical hiring task. In this, participants will be asked to express their hiring preference for a gender-neutral occupation between a man and a woman in one item, and a stereotypically masculine or a stereotypically feminine person in another. The same traits that are employed in the IRAP will be used in this task. A small number of self-report measures will also be included to assess whether the sample was relatively normative in terms of its explicit anti-women or gender-normative beliefs.

Method

Participants

Forty-seven undergraduate students (26 identified as female, 21 as male) aged between 18 and 42 years participated in this study (M_{age} = 23.84, SD = 5.49). The sample comprised all White Irish participants, with the exception of two White Western European individuals. Participation was voluntary and no remuneration was offered. Inclusion criteria included fluent English, normal or corrected-to-normal vision, and full use of both hands. Participants were provided with a fixed response format for gender and were given three options: female, male, or third.

Procedure

Experimental sequence. All experimental sessions were conducted one-to-one in individual experimental cubicles. Participants were briefed on the general nature and structure of the experiment and were given a short overview of the experiment's subject matter (i.e., contemporary beliefs about gender) prior to participation. Written informed consent was provided by the participant, followed by a verbal assessment by the researcher for all inclusion criteria. The general experimental sequence was as follows: hiring task, measures ascertaining self-reported beliefs towards women and gender, and two IRAPs. The order of the implicit measures and the presentation order of the blocks within them were both counterbalanced across participants. Upon completion of all tasks, participants were fully debriefed and thanked for their time.

Measures.

Self-report measures. To confirm whether the sample comprised relatively normative undergraduate students (i.e., not explicitly sexist or gender-conservative), participants also completed two short questionnaires. Attitudes toward women were ascertained by the

Modern Sexism Scale and towards more general gender-relevant matters by the Heteronormativity Attitudes and Beliefs Scale (HABS: Habarth, 2015).

Modern Sexism Scale. This 10-item scale, comprised of two five-item subscales, assesses beliefs about women and gender. The first subscale is intended to assess more traditional anti-women sentiments (e.g., "It is more important to encourage boys than to encourage girls to participate in athletics.") and the second more subtle or contemporary sexist attitudes (e.g., "Over the past few years, the government and news media have been showing more concern about the treatment of women than is warranted by women's actual experiences."). Items are scored on a Likert-scale from 1 (*strongly disagree*) to 5 (*strongly agree*), with possible scores ranging from 10-50. Higher scores indicate greater sexism.

Heteronormativity Attitudes and Beliefs Scale. The HABS is a 16-item questionnaire assessing heteronormative beliefs and assumptions. Heteronormativity can be broadly defined as the belief that people fall into one of two distinct gender categories (male and female), which form a natural heterosexual dyad. The HABs consists of two eight-item subscales assessing, first, "gender-as-binary" beliefs (e.g., "All people are either male or female") and, second, attitudes around natural or normative sexual behavior, such as the assumption of heterosexuality in men and women (e.g., "There are particular ways that men should act and particular ways that women should act in relationships"). Items are scored on a seven-point Likert scale from 1 (*strongly disagree*) to 2 (*strongly agree*), allowing a scoring range of 16-112. Higher scores indicate more pronounced heteronormative beliefs.

IRAPs.

Stimulus selection. Participants completed two gender binary IRAPs: the first contained positively valenced masculine and feminine traits, and the second contained negatively valenced traits. Stimuli for both IRAPs (Table 1) were obtained from a pilot study

(n = 234) in which a series of 60 personality traits were rated for both gender (i.e., masculinity-femininity) and desirability. Participants in the pilot study were aged between 18-25 years (117 female) and were undergraduate students at the same University as those in the current sample. Personality traits were ranked in terms of their combined gender and desirability ratings, and the four most desirable/undesirable masculine and feminine traits were selected for the positive and negative traits IRAPs, respectively.

Table 1 about here

Task structure. Prior to commencing the task, participants were provided with verbal instructions on how to complete the IRAP. These instructions broadly outlined the task structure (i.e., that they would be presented with blocks consisting of multiple word pairings and they would need to respond in accordance with a response "rule" presented before each block). The instructions also emphasized the importance of maintaining speed and accuracy throughout the task. Once participants were comfortable with these instructions, they began the "practice" phase of the IRAP, which was designed to train participants to a certain level of response fluency (78% accuracy and a median response latency of >2000 ms). Participants were presented with up to four pairs of practice blocks (i.e., four iterations of paired Rule A and Rule B blocks) until they reached the desired level of fluency, after which point they moved to the "test" portion of the IRAP. Those who did not meet the practice criteria did not complete the test blocks.

The practice and test phases of the IRAP were identical in terms of their stimuli and block structure. Both involved the presentation of a pre-block rule screen, 32 individual trials and a post-block feedback screen outlining the participant's accuracy and latency scores for that block. The pre-block rule screen presented Rule A or B (e.g., "Please respond as if men

have more stereotypically masculine traits and women more stereotypically feminine traits") and reiterated the task instructions:

> This task will determine what makes "intuitive sense" to you by seeing what rules you find easy and hard to follow. You'll pair words or images according to a rule. You'll be told the rule and when it changes. If you make a mistake, you'll see a red "X." Provide the correct response to continue. Learn to respond accurately according to the rule. When you've learned to be accurate you'll naturally speed up too. Going quickly without being accurate will not provide meaningful data.

On each trial, two stimuli were presented together on the screen (one label stimulus and one target stimulus, e.g., "women" and "nurturing"). The two response options (*true* and *false*) remained static across all trials at the bottom left and right of the screen. Each stimulus remained on the screen until the correct response was emitted. If participants responded incorrectly, corrective feedback in the form of a red "X" appeared in the center of the screen. Each block pair consisted of one "Rule A" block and one "Rule B" block. In the "Rule A" block, response contingencies reinforced choices of the on-screen word *true* when men-masculine and women-feminine stimulus pairings were present and *false* for men-feminine and women-masculine word pairings. In the "Rule B" block, the inverse response options were reinforced. The order of the rule blocks was always counterbalanced between participants.

Hiring task. Hiring preference was assessed using a brief task in which participants were presented with two identical questions: "If you were an employer hiring for an office job, which of the following two categories of people would you be more likely to hire?" The generic title "office job" was selected due to its non-specific nature and absence of any salient gender connotations (see Reuben, Sapienza, & Zingales, 2014 for a recent list of

gender employment stereotypes). Using the same traits as employed in the positive traits IRAP, the first question in this task ascertained hiring preferences for a stereotypically masculine person ("Someone who is witty, charismatic, competitive, and decisive") over a stereotypically feminine person ("Someone who is nurturing, gentle, affectionate, and sensitive"). The second item more explicitly asked for their preference for a man relative to a woman. For both items, participants were presented with a third response option: "I prefer not to answer." This was employed as a catchall for non-responses that may be due to any number of preferences (e.g., neutral/neither/both/disagree with the premise of the question, etc.) and to eliminate the possibility of inaccurate data produced by forced-choice responding.

Data processing and analysis

Following routine practices, latency differentials across Rule A and Rule B blocks were quantified using the D_{IRAP} scoring algorithm, a scoring metric based on an adaptation of Cohen's *d*. As previously mentioned, D_{IRAP} scores are analyzed at the trial-type level so as to provide an assessment of effect size for each individual trial-type (i.e., men-masculine, men-feminine, women-masculine, women-feminine). Thus, four separate D_{IRAP} scores were produced for each instance of the IRAP. D_{IRAP} scores range from +2 to -2, with positive D_{IRAP} scores representing a binary-consistent effect (e.g., responding to "men" and "masculine" with *true* faster than *false*), whereas negative D_{IRAP} scores represent binary-inconsistent biases (e.g., responding to "women" and "masculine" with *true* faster than *false*).

Practice-block data was not included in the analysis, and thus IRAP data was only collected from participants who progressed to the test phase (45 participants for the positive-traits IRAP and 44 for the negative-traits IRAP). Using the exclusion method outlined in Nicholson and Barnes-Holmes (2012), IRAP data were removed for participants who failed

to meet accuracy and/or latency criteria in more than one of the three test-block pairs. Three participants were removed from the negative traits IRAP on this basis. Participants were not excluded from the analysis if they failed to meet criteria in only one of their test-block pairs; however, the final D_{IRAP} scores for these individuals were calculated by averaging the D_{IRAP} scores across the remaining two (rather than three) pairs of test blocks. D_{IRAP} scores for three participants were calculated in this manner.

Results

All analyses were conducted using JASP (version 0.7.5 Beta 2, University of Amsterdam, Netherlands). All tests were two-tailed with alpha set at .05.

Hiring Preferences

When asked about their preference for a particular gender (i.e., male or female), responses in the current sample were varied (11% selected the man, 44% selected the woman, and 45% selected "I prefer not to answer"). A chi-square goodness-of-fit test revealed this distribution to be significantly unequal, χ^2 (2, n = 47) = 11.40, p = 0.003. For the stereotypical feminine/masculine-preference item, however, participants demonstrated an overwhelming preference for the masculine person (83% selected masculine traits, 13% selected feminine traits, and 4% selected "I prefer not to answer"). A chi-square goodness-of-fit test again revealed this to be a significantly unequal distribution χ^2 (2, n = 47) = 52.64, p < 0.001. Chi-square tests for independence revealed no significant differences between male and female responses for either item (all ps > .5).

Self-Report Measures

Self-reported sexism could be considered low to moderate in the current sample (females: M = 24.42, SD = 5.87; males: M = 28.9, SD = 7.43). Heteronormativity was also

relatively low although more varied than sexism scores (females: $M = 41.88$, $SD = 12.12$; males: $M = 50.4$, $SD = 18.11$). Independent samples t-tests identified significant gender differences on both the HABS, $t(46) = -1.9$, $p = .031$, and the MS, $t(46) = -2.26$, $p = .014$, with males scoring higher across the two scales. Effect sizes for both were large (Cohen's $d > .8$).

IRAP Performance

A visual inspection of the graph indicates that effects on all trial types reflected a binary-consistent pattern of responding (see Figure 2). Bias scores were typically larger for male participants, with the exception of the women-feminine trial-type. Larger resistance to forming role-incongruent relations was found for the men-feminine relative to the women-masculine trial type for both genders, with males demonstrating notably stronger "men-*not*-feminine" effects.

[Figure 2 about here]

A 2 × 2 × 4 repeated measures ANOVA was conducted to assess the impact of participant gender and IRAP type (i.e., positive or negative traits) on trial-type scores. While there was a significant two-way interaction between trial type and gender, $F(3, 35) = 3.94$, $p = .01$, $= .18$, no significant three-way interaction was found between trial-type, gender, and IRAP type, $F(3, 35) = 1.77$, $p = .16$, $= .046$. That is, although male and female participants performed significantly differently based on trial type within the IRAP, these effects were not related to the valence of the traits. There was no significant main effect for gender, although men did display marginally more binary-consistent biases than women on both IRAPs, $F(1, 37) = 3.4$, $p = .07$, $= .084$.

Follow-up Bonferroni-corrected independent-samples t-tests were conducted to explore the above two-way interaction effect between participant gender and IRAP trial-type (see Figure 2 for a graphical representation of trial-type scores). As performance on the IRAP did not differ according to trait valence, participant scores for both IRAPs were collapsed (at the trial-type level) for these analyses. Thus, only participants with data for both IRAPs were included here ($n = 40$). Differences were found on a single trial-type, the men-masculine trial-type: male participants demonstrated more positive D_{IRAP} scores ($M = 0.50$) than females ($M = 0.05$), $t(37) = 3.71$, $p < 0.001$. Men also demonstrated stronger "men-not-feminine" effects than women, though this was only a marginal effect ($M_{women} = .018$; $M_{men} = .159$), $t(37) = 1.95$, $p < 0.059$. This suggests that, across both of the male trial-types, men demonstrated stronger gender binary-consistent biases than women with regard to their own gender.

Measure Comparisons

Pearson's R correlations were conducted to explore the direction and significance of the relationship between explicit binarist or anti-women attitudes and IRAP performance (calculated using the overall D_{IRAP} scores). For male and female participants, no significant correlations were found between scores on the positive or negative IRAP and either the HABS or MS ($ps > .15$).

Results Summary

This study revealed significant binarist gender stereotypes in an undergraduate sample. Across two IRAPs, participants demonstrated effects in the expected role-congruent direction (i.e., men are masculine *and not feminine*, women are feminine *and not masculine*). Gender differences were identified in IRAP performances, with males demonstrating marginally larger gender stereotype biases across both IRAPs. Follow-up tests revealed this to be driven predominantly by differential performance on the "men-masculine" trial-type,

with males demonstrating significantly stronger effects on this trial-type. Hiring preferences were in the expected direction, with a significant proportion of participants preferring the masculine over the feminine candidate. Interestingly, when asked to express a preference for a male over a female candidate, nearly half of the participants responded that they would prefer not to answer, while the majority of the remaining sample elected the female. No significant correlations were found between scores on either IRAP and self-reported attitudes towards women and gender. Significant gender differences were found on the two self-report scales, with males demonstrating larger sexism and heteronormativity scores than females.

Discussion

In the current study, effects on both the positive and negative traits IRAPs were consistent with a binarist gender ideology. That is, participants readily (i.e., speedily) coordinated men *but not women* with stereotypically masculine traits, and women *but not men* with stereotypically feminine traits. While the absence of role-incongruent biases is relatively subtler than the strong "men-masculine" and "women-feminine" biases, both response patterns are theoretically important in suggesting that male and female traits may not merely be distinct, but also mutually exclusive. The ability to separate out specific biases (e.g., using the IRAP) therefore distinguishes the current work from previous studies (e.g., Cartwright et al., 2016; Rudman & Glick, 2001), and allows for stronger theoretical conclusions. For instance, the current study provides a starting point for investigating the potentially asymmetrical ways in which we "gender" men relative to women and the relative impact of participant gender on these behavioral patterns. In this study, IRAP effects were strongest for all participants on the trial type that was congruent with their own gender (i.e., the men-masculine trial type for males and women-feminine trial type for females). However, these effects were not symmetrical across males and females, with men demonstrating more

pronounced men-masculine and men-*not*-feminine biases than women on the comparable women-feminine and women-masculine trial types.

While the current findings require substantiation across different participant samples and methods of measurement, the results obtained here do cohere with existing theories of gender identity and gender role development. For example, several researchers have argued that masculinity is potentially a more rigid social construct than femininity with more well-defined boundaries (Bem, 1993; Leaper & Friedman, 2007; Thorne, 1993). More specifically, a number of observational studies have reported that although gender-appropriate behavior is typically rewarded in both girls and boys, displays of gender-nonconforming behavior tend to be more actively punished in boys (Adams & Coltrane, 2004; Bem, 1993; Kimmel & Messner, 2009; Leaper, 2002). The typical explanation for this effect is that many of our patriarchal or male-dominated social spheres (such as politics or business) place more value on masculine traits, meaning that gender-role deviations are more problematic for men than for women (see Coltrane & Adams, 2008). In more technical terms, if femininity is more undesirable for men then masculinity is for women, gender may participate in a more well established frame of opposition for men. However, this possibility requires examination using a measure more suited to the assessing the strength and nature of a single stimulus relation, rather than the concurrent operation of two sets of relata, as was done here using the IRAP.

The idea that greater cultural value is placed on masculine traits is evidenced by the hiring task responses observed in the current study. When asked which sort of person they would rather hire, participants expressed an overwhelming preference (83%) for the stereotypically "masculine" traits. While this question did not specify the sex of the candidate directly, the ascription of these traits to men but not women on the IRAP does suggest that the hirable traits were considerably more male than female. These data therefore support previous research identifying a link between masculine traits and competency or leadership

(Eagly & Carli, 2007; Koenig, Eagly, Mitchell, & Ristikari, 2011) and, again, the argument that Western societies continue to implicitly prioritize masculine values and traits (Broverman, Vogel, Broverman, Clarkson, & Rosenkrantz, 1972; Johnson, 2005; Serano, 2009).

It is interesting to note, however, that the responses on the direct gender-preference question (i.e., between a man and a woman) were not in a pro-male direction. A large proportion (45%) of the current sample elected not to express an explicit preference between male and female candidates (i.e., they selected the option "I prefer not to answer"); of those who did, however, the majority selected the female (44%) in favor of the male (11%). Though several factors may have contributed to this response pattern, the widespread unwillingness to express an explicit preference—at least a male one—is worth noting. For instance, it may reflect either self-presentational distortions (i.e., social desirability biases), which could attest to a growing awareness on behalf of participants of the issues women face in occupational contexts, or to a willingness to prioritize women, at least in theory.

More broadly, these data attest to the utility of focusing on the verbal construction of gender-as-binary in equality research, and of the use of the IRAP for this purpose. To date, much of the research into linguistic binaries and polarities has been conducted using more traditional sociological or feminist methods, such as questionnaires or interviews (e.g., Phipps, 2007). However, such approaches could be seen as potentially limited, given that they are notoriously time and labor-intensive and are susceptible to both introspection issues and social desirability biases (see Griffin & Phoenix, 1994; Nisbett & Wilson, 1977; Wood & Kroger, 2000). Further, although such approaches are undoubtedly rich and informative in extrapolating out the binary construction of gender, these often require themes or patterns to be abstracted from responses (e.g., via discourse analysis). In contrast, the IRAP allows for a

relatively quick, quantitative, and easy-to-administer assessment of linguistic binaries that is also conceptually coherent within an RFT approach to verbal behavior.

Future research should now seek to validate the binary IRAP using a larger and more representative sample. As mentioned previously, some theorists have argued that many forms of gender inequality and discrimination could be underpinned by assumptions of intrinsic male-female difference (e.g., Bem, 1993; see also Roche & Barnes, 1998). It would therefore be useful to have a validated measure of binarist beliefs for use in broader discrimination research (e.g., research into a propensity to sexually harass women). In addition, to assess the validity of the IRAP as a measure of gender attitudes outside the laboratory, further studies could compare performances between two known groups (e.g., gender fluidity advocates versus highly gender-conforming individuals), or the impact of relevant gender equality interventions (such as sexual consent workshops) on gender IRAP effects.

Some limitations to the current study should be noted. First, when providing information on their own gender, participants were provided with three options: male, female, and third. It is acknowledged that a more open-ended and in-depth means of collecting gender data is preferable to a box-ticking approach, given that it may have enabled an analysis of the role of individual gender identification in binarist views and/or gender discrimination. It could be the case, for example, that the men-masculine and women-feminine trial-types are differentially impacted by individual gender identity and/or conformance to stereotypical norms. Second, the current study treated the categories "men" and "women" as homogenous social groups rather than complex constructs that are intersected with class, race, ability, etc. (hooks, 1981; see also Orr, Taylor, Kahl, Earle, Rainwater, & McAlister, 2007). While it is difficult to bring that level of complexity to bear in an implicit measure, future research could perhaps explore this in more detail. For example, this could be done by employing additional pictorial IRAPs that contrast white men

and women with individuals from other racial or social groups, thereby examining these effects as they relate in turn to ethnicity.

Overall, findings from the current research provide additional support for the use of "conflicting relations" paradigms in assessing verbal histories around gender. In addition, the current findings attest to the potential conceptual benefit of hinging gender equality research on the verbal construction of gender-as-binary in language, especially as conceived in relational terms. The findings are consistent with existing theories suggesting that binarist stereotypes play a role in occupational discrimination against women, and broader arguments suggesting that gender could be more rigidly defined or constructed for males. Importantly, however, the current research approach allowed us to examine such potentially amorphous conceptions in technical and empirical terms, thereby bringing difficult-to-examine research topics into the behavior-analytic laboratory.

References

Adams, M. & Coltrane, S. (2004). Boys and men in families: The domestic production of gender, power, and privilege. In R. Connell, J. Hearn & M. Kimmel (Eds.). *The handbook of studies of men and masculinity.* Thousand Oaks, CA: Sage.

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.

Bem, S. L. (1993). *The lenses of gender: Transforming the debate on sexual inequality*. Yale University Press.

Broverman, I. K., Vogel, S. R., Broverman, D. M., Clarkson, F. E., & Rosenkrantz, P. S. (1972). Sex-role stereotypes: A current appraisal. *Journal of Social Issues*, *28*, 59–78. http://doi.org/10.1111/j.1540-4560.1972.tb00018.x

Butler, J. (1990). *Gender trouble: feminism and the subversion of identity*. Routledge.

Butler, J., & Butler, M. E. (2002). *Gender trouble: Tenth anniversary edition*. Routledge.

Cartwright, A., Roche, B., Gogarty, M., O'Reilly, A., & Stewart, I. (2016). Using a modified Function Acquisition Speed Test (FAST) for assessing implicit gender stereotypes. *The Psychological Record*, 1–11. http://doi.org/10.1007/s40732-016-0164-5

Coltrane, S. & Adams, M. (2008). *Gender and families* (2nd ed.). New York, NY: Rowman & Littlefield Publishers.

Connell, R. W., & Messerschmidt, J. W. (2005). Hegemonic masculinity: Rethinking the concept. *Gender & Society*, *19*, 829–859. http://doi.org/10.1177/0891243205278639

De Houwer, J., & Moors, A. (2010). Implicit measures: Similarities and differences. In B. Gawronski & K. B. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 176–193). New York, NY: Guildford Press.

Drake, C. E., Kellum, K. K., Wilson, K. G., Luoma, J. B., Weinstein, J. H., & Adams, C. H. (2010). Examining the Implicit Relational Assessment Procedure: Four preliminary studies. *The Psychological Record*, *60*, 81-100.

Eagly, A. H., & Carli, L. L. (2007). *Through the labyrinth: The truth about how women become leaders*. Boston, MA: Harvard Business School Press.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, *109*, 573–598. http://doi.org/10.1037/0033-295X.109.3.573

Gelman, S. A. (2005). *The essential child: Origins of essentialism in everyday thought*. New York, NY: Oxford University Press.

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

Grey, I. M., & Barnes, D. (1996). Stimulus equivalence and attitudes. *The Psychological Record*, *46*, 243–271.

Griffin, C., & Phoenix, A. (1994). The relationship between qualitative and quantitative research: Lessons from feminist psychology. *Journal of Community & Applied Social Psychology*, *4*, 287–298. http://doi.org/10.1002/casp.2450040408

Habarth, J. M. (2015). Development of the heteronormative attitudes and beliefs scale. *Psychology & Sexuality*, *6*, 166–188. http://doi.org/10.1080/19419899.2013.876444

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational Frame Theory: A post-Skinnerian account of human language and cognition*. New York, NY: Kluwer Academic/Plenum Publishers.

hooks, b. (1981). *Ain't I a woman: Black women and feminism*. Boston, MA: South End Press.

Hussey, I., Barnes-Holmes, D., & Barnes-Holmes, Y. (2015). From Relational Frame Theory to implicit attitudes and back again: clarifying the link between RFT and IRAP research. *Current Opinion in Psychology*, *2*, 11–15. http://doi.org/10.1016/j.copsyc.2014.12.009

Johnson, A. G. (2005). *The gender knot: Unraveling our patriarchal legacy*. Philadelphia, PA: Temple University Press.

Kessler, S., Ashenden, D. J., Connell, R. W., & Dowsett, G. W. (1985). Gender relations in secondary schooling. *Sociology of Education*, *58*, 34–48. http://doi.org/10.2307/2112539

Kimmel, M. S. (2000). *The gendered society.* New York, NY: Oxford University Press.

Kimmel, M. S. (2007). *The sexual self: The construction of sexual scripts*. Nashville, TN: Vanderbilt University Press.

Kimmel, M. S. & Messner M. A. (2009). *Men's Lives* (7th Ed.). Boston, MA: Allyn & Bacon.

Koenig, A. M., Eagly, A. H., Mitchell, A. A., & Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin*, *137*, 616–642. http://doi.org/10.1037/a0023557

Leaper, C. (2002). Parenting girls and boys. In M. H. Bornstein (Ed.), *Handbook of parenting: Vol. 1. Children and parenting.* (2nd ed., pp. 189 – 225). Mahwah, NJ: Erlbaum.

Leaper, C., & Friedman, C. K. (2007). The socialization of gender. In J. Grusec & P. Hastings (Eds.), *Handbook of socialization: Theory and research* (pp. 561 – 587). New York: Guilford.

Moxon, P., Keenan, M., & Hine, L. (1993). Gender-role stereotyping and stimulus equivalence. *The Psychological Record*, *43*, 381–394.

Nicholson, E., & Barnes-Holmes, D. (2012). The Implicit Relational Assessment Procedure (IRAP) as a measure of spider fear. *The Psychological Record*, *62*, 263–278.

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

O'Reilly, A., Roche, B., & Cartwright, A. (2014). Function over form: A modern behaviour-analytic approach to implicit attitudes. In J. Zin (Ed.), *Exploring implicit cognition: Learning, memory, and social-cognitive processes.* Philadelphia, PA: IGI Global.

O'Reilly, A., Roche, B., Ruiz, M., Tyndall, I., & Gavin, A. (2012). The Function Acquisition Speed Test (FAST): A behavior analytic implicit test for assessing stimulus relations. *The Psychological Record*, *62*, 507–528.

Orr, D., Taylor, D., Kahl, E., Earle, K., Rainwater, C. & McAlister, L. (Eds.) (2007). *Feminist politics: Identity, difference, and agency*. Lanham, MD: Rowman & Littlefield.

Phipps, A. (2007). Re-inscribing gender binaries: Deconstructing the dominant discourse around women's equality in science, engineering, and technology. *The Sociological Review*, *55*, 768–787. http://doi.org/10.1111/j.1467-954X.2007.00744.x

Rabelo, L. Z., Bortoloti, R., & Souza, D. H. (2014). Dolls are for girls and not for boys: Evaluating the appropriateness of the Implicit Relational Assessment Procedure for school-age children. *The Psychological Record*, *64*, 71–77. http://doi.org/10.1007/s40732-014-0006-2

Reuben, E., Sapienza, P., & Zingales, L. (2014). How stereotypes impair women's careers in science. *Proceedings of the National Academy of Sciences*, *111*, 4403–4408. http://doi.org/10.1073/pnas.1314788111

Roche, B., & Barnes, D. (1998). The experimental analysis of human sexual arousal: Some recent developments. *The Behavior Analyst*, *21*, 37.

Roche, B., Barnes-Holmes, Y., Barnes-Holmes, D., Stewart, I., & O'Hora, D. (2002). Relational Frame Theory: A new paradigm for the analysis of social behavior. *The Behavior Analyst*, *25*, 75–91.

Roche, B., & Dymond, S. (2013). *Advances in Relational Frame Theory: Research and application*. Oakland, CA: New Harbinger Publications.

Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues*, *57*, 743–762.

Serano, J. (2009). *Whipping girl: A transsexual woman on sexism and the scapegoating of femininity*. Berkeley, CA: Seal Press.

Thorne, B. (1993). *Gender play: Girls and boys in school.* Buckingham, UK: Open University Press.

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

Witt, S. D. (1997). Parental influence on children's socialization to gender roles. *Adolescence*, *32*, 253–259.

Wood, L. A., & Kroger, D. R. O. (2000). *Doing discourse analysis: Methods for studying action in talk and text*. Thousand Oaks, CA: Sage Publications.

Compliance with Ethical Standards

Funding: This research was funded by a Maynooth University John and Pat Hume research scholarship awarded to the first author.

Conflict of Interest: The Authors declare that there is no conflict of interest.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.