



Technical Report

Prepared by:
The Centre for Ethical Leadership
Ormond College
The University of Melbourne

Research team:

Michelle Stratemeyer (Research Manager)

Dr Victor Sojo

Dr Melissa Wheeler

Dr Vanja Rozenblat

Dr Ilro Lee

Dr Divya Peter

Marina Kociski

Melanie McGrath

Anna Genat

Professor Robert Wood (Research Director)

The authors acknowledge the support of Cognicity Pty Ltd who provided the Cognicity Academy on-line training program and Cognicity Laboratories unconscious knowledge measures used in the test of unconscious bias training.

The authors acknowledge the contribution of Chapter 2 to this Report: Intervention 2: Debiasing job advertising by the Victorian Health Promotion Foundation (VicHealth) Leading Thinkers: Behavioural Insights and Gender Equality initiative.

**CENTRE FOR
ETHICAL
LEADERSHIP**



Contents

| | |
|---|----|
| Executive Summary | 1 |
| Intervention 1: Targeted Recruitment | 4 |
| Case Study: TAC | 6 |
| Intervention 2: De-biasing Language in Job Advertisements (VicHealth) | 11 |
| Intervention 3: CV De-Identification | 19 |
| Case Study 1: VicRoads | 22 |
| Case Study 2: Hall & Wilcox | 27 |
| Case Study 3: Department of Premier and Cabinet | 32 |
| Case Study 4: Department of Treasury and Finance | 37 |
| Case Study 5: Department of Justice and Regulation | 42 |
| Intervention 4: Unconscious Bias Training | 48 |
| Case Study: Unconscious Bias Training across Victorian Government Departments | 51 |
| Full Reference List | 57 |
| Appendix A | 60 |
| Appendix B | 62 |

Executive Summary

Recruit Smarter trialled four pilot interventions across a range of government departments and private sector organisations. The four interventions included two trials of targeted recruitment via modified language use in job advertisements, a CV de-identification program, and the provision of training to address unconscious bias. The pilot program has found evidence that these interventions are beneficial to improving equity of opportunity for diverse Victorian applicants.

Summary of results

First, a trial of targeted recruitment language in job advertisements, conducted by the Transport Accident Commission, found that increasing diversity-supportive language and inviting applications from candidates with disabilities improved the proportion of applicants who requested reasonable adjustments. Additionally, applicants with reasonable adjustment needs progressed further through recruitment after the introduction of the advertisement, suggesting that simply increasing the number of applicants from a minority group may assist progress through recruitment. This should be supported by further strategies, such as unconscious bias training or disability-friendly workplace policies, which may also improve the progression of applicants with disabilities through recruitment processes.

Second, five organisations trialled de-identification of demographic information from applicant CVs during recruitment. These trials were conducted across a range of organisations and roles. Results showed that CV de-identification was successful in improving equity for applicants in specific organisations; gender equity was improved at the Department of Treasury and Finance, socioeconomic status equity was improved at the Department of Premier and Cabinet, and overseas-born applicants

moved from a position of disadvantage to advantage at VicRoads. However, some characteristics were not significantly affected by de-identification; this was especially the case when there was little discrepancy between different social groups in the baseline recruitment process, or in cases when minority groups were not well represented in the applicant pool to begin with. These non-significant results suggest that de-identification should be carefully implemented for specific roles that suffer from inequity in hiring across social groups. Organisations that already have good diversity in candidates throughout the recruitment process – from the application to the offer stage – may instead focus on other aspects of recruitment, such as ensuring that induction and onboarding processes are comprehensive and assist in integrating new staff members, to ensure that diverse staff are retained in the longer term.

Third, a trial of unconscious bias training across seven organisations found that training improved perceptions of self-efficacy for diversity-supportive behaviour and behavioural intentions to support diversity, which lead to an increase in diversity-supportive behaviours back on the job; for example challenging a colleague who makes a biased comment or initiating conversations about diversity in the workplace. These findings are in accordance with previous research that shows beneficial results from unconscious bias and diversity training (Bezrukova, Spell, Perry, & Jehn, 2016). Though the improvements from this training are relatively small, these can accumulate over time, especially if a large proportion of staff members in an organisation are provided with training.

Organisations should carefully consider the content and context of unconscious bias training initiatives. Some research suggests that mandatory training may result in stronger positive effects of training (e.g. Ellis & Sonnefield, 1994). This may be due to self-selection effects; voluntary training is more likely to engage staff who already support diversity initiatives, whereas mandatory training includes staff that do not necessarily hold diversity-supportive attitudes. However, Ellis and Sonnefield caution that mandatory training may be resented by staff unless the corporate culture is one which encourages company-wide policies and uniformity in training. Other considerations include the length and content of training programs, with evidence supporting lengthier, integrated programs (Bezrukova et al., 2016) that include specific, rather than general strategies and scenarios applicable to the workplace (Emerson, 2017).

Broadly speaking, the interventions described in this report have shown positive results for equity of access for diverse job applicants. To support and maintain these positive changes, these types of interventions must be supplemented by more general policies and practices that support diverse workplace environments. Integrated approaches to diversity and inclusion tend to work better than a scattergun approach, as the organisation signals its support of diversity, and different programs of work reinforce each other's learning points. Leadership is also key to successful diversity and inclusion initiatives (Jayne & Dipboye, 2004).

Organisational leadership and championing of diversity is critical for creating a visibly non-discriminatory workplace environment that offers staff the opportunity to develop their careers and reach their potential. Organisations can consider strategies such as including diversity and inclusion initiatives as key performance indicators for managerial staff, to hold managers accountable for reflecting the broader population in the organisation's staffing (Jayne & Dipboye 2004).

Future research directions

There are a number of directions that future research programs could consider as follow-ups to the Recruit Smarter pilot program.

First, most of the roles included in this program were not managerial roles but instead clustered in the VPS4 to VPS6 range for government organisations. Diversity of staff tends to decrease for roles that involve increasing levels of leadership.

For example, women tend to become severely underrepresented at higher levels of leadership, and cultural diversity also suffers through the so-called 'Bamboo Ceiling'— an underrepresentation of leaders from Asian cultural heritage (Australian Human Rights Commission, 2018).

These trends conflict with the research demonstrating the greater effectiveness of diverse leadership teams. For example, leadership styles typically ascribed to women have been found to positively influence working environments. A meta-analysis by Eagly, Johannsen-Schmidt and Engen (2003) found that women were more likely than men to deliver appropriate feedback and reward behaviours – characteristics that are predictive of effective performance. (Lowe, Kroeck, & Sivasubramaniam, 1996). Similarly, leaders from ethnic minority backgrounds are also more likely to engage in transformational leadership, characterised by inspiring, respectful, and authentic practices (Ardichvili, Mitchell, & Jondle, 2009).

Such research suggests the importance of equity of access to leadership roles for diverse candidates. Future research programs may wish to target managerial and executive roles for interventions to determine the best ways of ensuring equality of access to these roles.

While ensuring a diverse range of applicants in leadership roles is critical, organisations should also be aware of phenomena that may undermine leaders drawn from underrepresented groups. One such example is the 'glass cliff' phenomena (Ryan & Haslam, 2005). The glass cliff refers to cases where a woman is appointed as a leader of an organisation during times of crisis. These leadership roles are often precarious and volatile, which may result in female leaders facing additional hurdles to successful management of an organisation. Organisations should be aware of such research, and the potential undermining of diverse leaders in the types of roles that are accessible to them.

Finally, organisations are cautioned that the results of this pilot program predominantly focus on progression through recruitment (e.g., from application, shortlisting, interview, through to hiring). However, the success of these programs ultimately depends on hiring outcomes—that is, whether diverse applicants are offered roles, whether those applicants accept the roles, and whether organisations can retain their diverse talent over time. Minority applicants are not necessarily benefited by progressing further through recruitment without any discernible difference in job offers.

Therefore, it is critical that organisations audit their staff diversity on a regular basis to determine whether recruitment strategies and policies are resulting in diversity of staff across roles, seniorities, and divisions of labour.

Intervention

1

Targeted Recruitment

Targeted recruitment refers to a set of strategies undertaken to increase applicants and new hires from specific groups, typically those under-represented in the organisation.

Strategies tend to involve a focus on targeted communication, such as engaging with advocacy groups and advertising in community media. In particular, altering job advertisement language is considered to be a particularly effective strategy for targeted recruitment.

Steps to implement

Changing the language of job advertisements can ensure greater attractiveness to diverse candidates. This is an easy and inexpensive strategy for increasing the number of minority applicants and broadening the skilled worker labour pool available to the recruiting organisation (Araten-Bergman, 2016).

Key features include:

- Including phrases specifically noting that people from the target minority group are welcome to apply. For example “applications from members of [minority group] are encouraged.”
- Including a separate contact name or phone number for applicants with specific requirements (e.g., those living with a disability).
- Providing a general statement about values and commitments to diversity and inclusion, for example “we are an equal opportunities employer” or “we welcome candidates from diverse backgrounds.”
- When recruitment involves visual or audio features, including images of people from diverse backgrounds, including people with disabilities. Television and radio advertising should include a variety of voices, including males and females and people with accents.

Measuring outcomes

After creating two versions of the advertisement – the original, standard version and the modified diversity-supportive version, the proportion of minority applicants is compared. A significant increase in minority applicants following the introduction of the modified advertisement would indicate the effectiveness of the job advertisement. Baseline and post experimental trials intervention comparisons (or pre-test/post-test design) further strengthen confidence in conclusions drawn from the results of the study.

Case Study on Targeted Recruitment

TAC

The Transport Accident Commission (TAC) provides support to individuals injured on Victorian roads, many of whom may acquire a disability as a result of their injuries. In line with this mission, TAC's targeted recruitment intervention aimed to increase the number of applications with disabilities applying for positions within the organisation.

Method

Participants

The sample for this dataset comprised 3341 applicants (1834 female, 1331 male, 176 unspecified) who applied for one of 149 roles within the TAC from July to October 2016 (baseline) or July to October 2017 (experimental). The baseline condition comprised 1744 applicants, with 1597 applicants in the experimental condition. Demographic variables are summarised in Table 1. Complete data for the reasonable adjustments question was obtained for $n = 2637$ participants.

Study Design

The study included measures of applicant responses and recruitment outcomes prior to and following the introduction of the modified recruitment advertisement. Study participants responded to either the standard advertisement (Figure 1) or the modified advertisement that encouraged people with disabilities to apply (Figure 2).

Procedures and Measures

TAC's aim was to increase recruitment of individuals with a disability within their organisation. Accordingly, the intervention was a modification of the diversity statement in their standard advertisement to include a statement encouraging people with disabilities to apply and to identify themselves as requiring reasonable adjustments.

During the baseline period from July to October 2016, position advertisements contained the regular statement supporting diversity provided in Figure 1.

During the experimental period, from July to October 2017, the modified statement shown in Figure 2 was included in the advertisement.

About The TAC

The Transport Accident Commission (TAC) is a Victorian Government-owned organisation whose role is to promote the Towards Zero vision, improve the State's trauma system and support those who have been injured on our roads.

The Transport Accident Commission promotes a workplace that actively seeks to include, welcome and value unique contributions of all people. People from indigenous, culturally and linguistically diverse backgrounds, and people with disabilities are encouraged to apply.

If you require any adjustments to assist you with your application, please contact [name] on (xx) xxxx xxxx or alternatively send an email to xxx@tac.vic.gov.au.

The key outcome variable that was compared between the baseline and experimental conditions was the number of applicants self-identifying as requiring 'reasonable adjustments' in the workplace in the event they were employed. 'Reasonable adjustment' typically represents a modification to the workplace or work schedule required to accommodate a physical or mental disability. This unit of measure identifies candidates who are comfortable disclosing a disability during the recruitment process but does not include applicants who have a disability but do not wish to request or disclose a need for reasonable adjustments.

Figure 1. Original text appearing in TAC employment advertisements from July 1 to Oct 5, 2016.

About The TAC

The Transport Accident Commission (TAC) is a Victorian Government-owned organisation whose role is to promote the Towards Zero vision, improve the State's trauma system and support those who have been injured on our roads.

The TAC promotes a workplace that actively seeks to include, welcome and value unique contributions of all people and is committed to diversity and social inclusion in its employment practices.

We encourage applications from people with a disability as Victoria's commission for transport accident treatment and services; TAC is committed to assisting people with disabilities to achieve their career aspirations. People from indigenous, culturally and linguistically diverse backgrounds are also encouraged to apply. The TAC contributes to a fairer playing field for all Victorian job applicants, including making reasonable adjustments to the recruitment and employment process.

If you have specific requirements we need to consider, for example an adjustment to the recruitment process, please contact [name] on (xx) xxxx.xxxx or email xxx@tac.vic.gov.au.

Figure 2. Modified text appearing in TAC employment advertisements from July 1 to October 5, 2017.

Analysis

Chi-squared analysis was used to compare the proportions of individuals with and without a disability who applied for a role at the TAC across the two conditions (baseline and experimental). T-test were used to explore the difference in how far individuals with and without a reported disability progressed through the recruitment process. Two separate One-Way Analyses of Variance (ANOVA) models were then used to compare recruitment progress between the three groups at each of the two time points. The three research questions were:

1. Does the proportion of people with disabilities applying to TAC increase following the introduction of the disability-targeted advertisement?
2. Do applicants who indicate a need for reasonable adjustments progress further through recruitment following the introduction of the disability-targeted advertisement?
3. Are there differences in progress through the recruitment process between those who identify as having a disability, those who do not identify as having a disability, and those who declined to respond to the question?

Results

Demographics

Demographic data for applicants in the baseline and experimental conditions are shown in Table 1.

Table 1
Gender, Job Rank, and Reasonable Adjustment Requests by Condition

| | | Baseline Number (%) | Experimental Number (%) |
|------------------------|-------------|---------------------|-------------------------|
| Gender | Female | 943 (58.2%) | 891 (57.6%) |
| | Male | 676 (41.8%) | 655 (42.4%) |
| Job rank | Low | 368 (21.1%) | 508 (31.8%) |
| | Moderate | 858 (49.2%) | 748 (46.8%) |
| | Manager | 446 (25.6%) | 293 (18.3%) |
| | Executive | 72 (4.1%) | 48 (3.0%) |
| Reasonable adjustments | No | 1311 (75.2%) | 1268 (79.4%) |
| | Yes | 17 (1.0%) | 41 (2.6%) |
| | No response | 416 (23.9%) | 288 (18.0%) |

Research Question 1

Does the proportion of people with disabilities applying to TAC increase following the introduction of the disability-targeted advertisement?

A greater proportion of applicants indicated a need for reasonable adjustments when applying for roles at TAC after the introduction of the disability-targeted job advertisement. Before the intervention, 17 applicants (1% of all applicants) indicated a need for reasonable adjustments. After the intervention, 41 applicants (2.6% of all applicants) indicated a need for reasonable adjustments. There was also a decrease in the proportion of candidates who did not answer the question, from 23.9% prior to the intervention, to 18% after the introduction of the modified advertisement.

The increase in the proportion of applicants identifying as requiring reasonable adjustment was statistically significant, $\beta = 27.51, <.001$, and was matched by a decrease in applicants who did not respond to the question.

However, there were no major changes to the expected number of applicants who did not need reasonable adjustments in the pre- or post-intervention groups.

Research Question 2

Do applicants who indicate a need for reasonable adjustments progress further through the recruitment process following the introduction of the disability-targeted advertisement?

Applicants who indicated a need for reasonable adjustments progressed further through the recruitment process after the introduction of the new job advertisement ($M = 1.625, SD = 1.148$), compared to before the advertisement introduction ($M = 1.177, SD = .529$). This was a significant difference, $t(54.53) = -2.02, p = .048$.

Applicants who did not indicate a need for reasonable adjustment showed no difference in how far they progressed through the recruitment process after the introduction of the new job advertisement ($M = 1.670, SD = 1.333$), compared to before the advertisement introduction ($M = 1.734, SD = 1.528$). This was not a significant difference, $t(2419.81) = 1.11, p = .268$.

Finally, applicants who did not respond to the question regarding whether they did or did not have a disability showed a reduced progression through the recruitment process after the introduction of the new job advertisement ($M = 2.385, SD = 1.528$) compared to during the baseline ($M = 2.777, SD = 2.075$). This was a significant difference, $t(653) = 2.45, p = .015$. These differences are shown in Figure 3.



Figure 3. Average progress of applicants through decision points (bars indicate standard error) in the recruitment process for the standard and modified job advertisements, grouped by response to the reasonable adjustments question. Decision points were 1 = Unsuccessful; 2 = Reviewed by hiring manager; 3 = Interview 1; 4 = Interview 2. Three decision points were not included in the graph: 5 = Reference check; 6 = Talent pool; 7 = Offered a role.

Research Question 3

Are there differences in recruitment progress between those who identify as having a disability, those who do not identify as having a disability, and those who declined to respond to the question?

At the baseline, a one-way ANOVA suggested a significant difference in recruitment progress between the three groups, $F(2, 1630) = 59.29, p < .001$. Planned comparisons showed that those who reported no disability progressed further through recruitment than those reporting a disability, $t(19.90) = 4.12, p = .001$. Furthermore, those who did not respond to the question showed further progression through the recruitment process than those who either disclosed a disability or disclosed having no disability, $t(179.93) = -10.52, p < .001$.

In the experiment condition with the targeted advertisement, a one-way ANOVA again revealed a significant difference in recruitment progress between the three groups, $F(2, 1549) = 27.21, p < .001$. However, this time planned comparisons revealed no difference between how far candidates who disclosed a disability and those who disclosed no disability progressed through the recruitment process, $t(42.46) = .25, p = .808$. Those who did not respond to the question again showed further progression through the recruitment process than those who either disclosed a disability or disclosed having no disability, $t(204.91) = -4.93, p < .001$.

Discussion

The results show that targeted recruitment of people with a disability through an advertisement encouraging them to apply had a positive impact on both the proportion of applicants and their progression through the recruitment process. A simple change to the language of the advertisement produced an increase in the proportion (and number) of applicants who indicated they required reasonable adjustments. This was paired with a decrease in the number of applicants who did not respond to the question, which may indicate that applicants with a disability felt more comfortable disclosing their need for reasonable adjustments after the introduction of the new advertisement.

A secondary analysis demonstrated a further effect of the intervention. Applicants who indicated that they required reasonable adjustments progressed further through the recruitment process after the introduction of the new job advertisement language, while applicants who did not require reasonable adjustments were equally likely to progress before and after the advertisement introduction. This suggests that the benefits of the advertisement persist beyond attracting a wider pool of talent into later stages of recruitment. The reasons for this effect cannot be established conclusively from the TAC data. One possible explanation is that the targeted advertisement attracted applications from more capable people with disability, who might otherwise not seek employment with an organisation that they believe will not utilize their capabilities. A second possible explanation is that those conducting the interviews were more alert to their potential for bias against people with disabilities and took steps to mitigate the potential effects of that bias.

Finally, a third analysis showed for those who responded to the standard advertisement, applicants who did not disclose a disability progressed further through the hiring process compared to applicants who disclosed a disability. This gap decreased and was no longer significant after the introduction of the modified advertisement. For both advertisements, applicants who did not identify their disability status progressed further than those reporting either a disability or no disability. It is difficult to interpret this finding without more information on why applicants choose not to respond to the question. Interviews could be used to ascertain applicants' reasons for not disclosing their disability status, and hiring managers' perceptions and opinions of those who do not respond.

Intervention

2

De-biasing Language in Job Advertisements

This chapter and research contributed to the Recruit Smarter initiative by the Victorian Health Promotion Foundation (VicHealth) Leading Thinkers: Behavioural Insights and Gender Equality initiative.

Does changing the wording of a job advertisement change the gender representation of the applicant pool?

Summary

If asked, most of us would say that discrimination against women in the workplace is wrong and unacceptable. Yet research^{12,3} shows that treating women and men equally in hiring decisions, job evaluations and leadership positions remains far more of an ideal than a reality.

VicHealth's *Leading Thinkers: Behavioural Insights and Gender Equality* initiative aligned strongly with Recruit Smarter's collective approach to address unconscious bias in recruitment. Our behavioural insights trial with workplaces examining gendered language in job advertisements, was an opportunity to investigate the recruitment process as a potential key structural influencer of gender equality practice. VicHealth's overall initiative aimed to support the development of an evidence base on behavioural strategies that will advance gender equality across a range of settings such as workplaces and sports.

Harvard University Professor Iris Bohnet, VicHealth's Leading Thinker and author of *What Works: Gender Equality by Design* advocates for de-biasing organisations instead of individuals to address persistent unconscious bias. Professor Bohnet, a leading global expert in behavioural insights and gender equality⁴, led the design and methodology of this trial and advised VicHealth staff on key aspects of the trial during the delivery.

The recruitment process relies on human decisions, and evidence from the behavioural sciences show that the way we make decisions can be biased. We may be affected by unconscious biases during the process of attracting, selecting and promoting employees.

Key insights

The trial scoping process involved face-to-face interviews with approximately 25 organisations to determine organisational readiness for the trial, and a systems and process review. Following this, a number of organisations provided the first sweep of job advertisements for debiasing. We were able to surmise several top strategic insights from the interviews and examining the ads, including:

- There was already a high degree of gender neutrality in the language used in the adverts developed by our partner organisations.
- There is an increasing uptake by organisations to utilise the services of companies that address unconscious bias in recruitment such as Textio and Applied. We found that some of the organisations we spoke to used gender decoders but had not systematically evaluated the impact the decoders had on their recruitment practices.
- Across the board, partners identified the biggest problem they faced in their workforces was the lack of women in senior roles. Interestingly, awareness of the problem was not enough to address it, emulating the often-noted gap between intention and behaviour, found in behavioural sciences literature.⁵ (See Box 1)

There was not sufficient data collected in this trial, within the trial parameters, to claim that gendered language affects the gendered make-up of the applicant pool. However, there is an increasing body of evidence to suggest this is the case.

The Victorian Government's focus on gender equality has meant that many employers are focusing on strategies to make workplaces gender equitable. Focusing on gendered language in job advertisements was attractive to our trial partners, as it was at no cost to their organisation and appeared a relatively simple process of data collection.

1: Workplace Gender Equality Agency 2016a, Australia's gender equality scorecard: Key findings from the Workplace Gender Equality Agency's 2015–16 reporting data, WGEA, p. 9, viewed 22 July 2018, www.wgea.gov.au/sites/default/files/80653_2015-16-gender-equality-scorecard.pdf.

2: Azmat, G & Petrongolo, B 2014, 'Gender and the labor market: What have we learned from field and lab experiments?', *Labour Economics*, vol. 30, pp. 32–40.

3: Booth, A & Leigh, A 2010, 'Do employers discriminate by gender? A field experiment in female-dominated occupations', *Economic Letters*, vol. 107, no. 2, pp. 236–8.

4: Iris Bohnet, the Roy E. Larsen Professor of Public Policy, is the Academic Dean of Harvard Kennedy School. She is a behavioural economist, combining insights from economics and psychology to improve decision-making in organisations and society, often with a gender or cross-cultural perspective. Her most recent research examines behavioural design to de-bias how we live, learn and work.

5: Sheeran, P. & Webb, T.L. (2016, September) *The Intention–Behavior Gap*. *Social Personality Psychology Compass*, Volume 10, Issue 9, Pages 503–518.

Principles to remember

By Professor Iris Bohnet,
VicHealth's Leading Thinker and author of
What Works: Gender Equality by Design

Do:

- Experiment with the wording of job listings by removing adjectives closely associated with a particular gender. Software programs that highlight stereotypically gendered words can help counteract the candidate's perception that they do not belong in the work environment.
- Ask candidates to take a work sample test — it's useful in comparing applicants and it's an effective predictor of future job performance. Skills tests force employers to critique the quality of a candidate's work versus unconsciously judging them based on appearance, gender, age, and personality.
- Control for your personal feelings about a particular candidate by giving likability a numerical score. If it matters to you whether you like the person you hire then add likeability to a selection matrix. By giving likability a score, you're making it more controllable.

Don't:

- Engage in unstructured interviews. Instead, standardise the interview process by asking candidates the same set of defined questions. This minimises bias by allowing employers to focus on the factors that have a direct impact on performance.
- Allow surface demographic characteristics to play into your résumé review. Use a software program that blinds that information and ensures a level playing field.
- Neglect to set diversity goals. Be sure to track how well you're doing on them. They can be polarising for traditionally advantaged groups, so ensure you communicate the positive business advantages that diversity can bring.

De-biasing job advertisements – a trial in the Recruit Smarter pilot

Australia has historically had very high rates of occupational sex segregation – indeed in the mid-1980s, it had the most sex-segregated labour force in the OECD.⁶ This has persisted over the last two decades. Segregation has remained fairly constant across most occupations over time, with women making some inroads into management and professional occupations, but not into other male dominated occupational categories such as technicians. It is especially marked for personal assistants and secretaries (98 per cent female), and carpenters and joiners (0.7 per cent female).⁷

The extent of segregation is such that it has been estimated over half of Australian women would have to change occupations in order to have the same occupational distribution as Australian men.⁸

The VicHealth trial, *De-biasing Language in Job Advertisements*, was designed to examine unconscious bias in the first stage of recruitment—the job advertisement—and trials the effect of de-biased language on the gender composition of the applicant pool.

If simple differences in how a job is described and advertised is able to influence the gender ratio of the applicant pool, this will have important implications for a large range of employers in the public and private sectors seeking gender balance in their workforces. If it is shown that using readily available software and taking care to avoid specific types of words and descriptions is effective, private firms and government employers can go a long way towards equalising the pipeline of applicants for jobs traditionally dominated by men.

Testing this hypothesis in a 'real world' setting, VicHealth worked with a number of trial partners from government and industry, who are committed to fairer recruitment, to ultimately diversify their workforces. It was proposed that the trial findings will help employers determine where to focus their energy. If gendered language is shown to have an effect on the applicant pool, then employers can use gender decoders more readily. If the findings show no effect, conversely employers can focus their efforts elsewhere.

6: Moskos, M 2012, 'How occupational sex segregation shapes low-skilled men's employment opportunities in Australia', *Labour & Industry: A journal of the social and economic relations of work*, vol. 22, no. 4, pp. 415–32.

7: Australian Bureau of Statistics 2011.

8: Coelli, MB 2014, 'Occupational differences and the Australian gender wage gap', *Australian Economic Review*, vol. 47, no. 1, pp. 44–62.

There is a stream of research analysing the gendered nature of language.^{9, 10, 11, 12, 13} More specifically, this study draws heavily on research by Gaucher et al. 2011¹⁴ as it further directly relates to the question at hand. Gaucher et al. set out to measure how gendered the wording in job advertisements was, and whether the use of words associated with stereotypical gender roles, such as competence for men and warmth for women, has an impact on how people perceive these jobs.

In laboratory experiments examining the impact of masculine wording, authors found that people inferred from the ads how male or female-dominated the profession was. The more women inferred a profession to be male-dominated, the less appealing they found the job. Tellingly, it was not a matter of perceived competence to succeed at the job. Gendered wording told the experimental subjects something about whether or not they 'belonged' but did not affect whether or not they thought they had the skills to do the job.

The study indicated that subtle differences in the wording used to describe a job can affect the makeup of a candidate pool. Other research^{15, 16, 17} has also shown that the use of traditionally gendered language can influence the gender split of those who apply for a job.

The presence of this wording in advertisements and position descriptions may be sufficient to apply unconscious and often significant downstream consequences on each individual's appraisal of the relevant jobs.

Methodology

The research methodology consisted of two versions of the job advertisement: a control version and an intervention (treatment) version. The control version was the organisation's usual advertisement which may contain gendered language. The treatment version was a gender-neutral version that had undergone de-biasing.

Ideally this study is best suited to a Randomised Control Trial (RCT) design, where both advertisements, the control and treatment, are posted in the field at the same time. This design would offer the least amount of trial 'noise' – important when many of the ads had so few changes to the treatment.

While the trial was deemed 'low risk' in the ethics application, it was noted by the Ethics Committee that an RCT, with both advertisements in the field at the same time, could influence the trajectory of a candidate's career, and requested that this design be abandoned. To progress the trial, it was decided to offer an alternative trial design: a 'difference-in difference' model, where data was collected on previous listings of the same job advertisement at several time points in the past, as well as data collected on the treatment. The response to the treatment version would be compared to advertisement responses received to the organisation's previously used control version. See Figure 4.

To enhance the scientific validity of the trial, organisations were also asked to provide a similar advertisement to the one that was debiased – a matched advertisement – and these ads remain unchanged. This allowed the researchers to not only compare before and after, but it held the time dimension constant.

9: Bem, S. L. (1974) The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, pp. 42, 155-162.

10: Colley, A., G. Mulhern, J. Maltby, A.M. Wood. (2009). The short form BSRI: Instrumentality, expressiveness and gender associations among a United Kingdom sample. *Personality and Individual Differences*, pp. 46, 384-387.

11: Pedhazur, E. (1979). Bem Sex Role Inventory: A Theoretical and Methodological Critique. *Journal of Personality and Social Psychology*, pp. 996-1016.

12: Carver, A.F., A. Vafaei, R. Guerra, A. Freire, and S. P. Phillips (2013). Gender Differences: Examination of the 12-Item Bem Sex Role Inventory (BSRI-12) in an Older Brazilian Population. *Plos ONE*, 8, e76356.

13: Hoffman, R.M. & Borders, L.D. (2001). Twenty five years after the Bem Sex Role Inventory: A reassessment and new issues regarding classification variability. *Measurement and Evaluation in Counseling and Development*, pp. 34, 39-55.

14: Gaucher, D., Friesen, J., & Kay, A. (2011) Evidence That Gendered Wording in Job Advertisements Exists and Sustains Gender Inequality. *Journal of Personality and Social Psychology* Vol. 101, No. 1, pp. 109-128.

15: Formanowicz, M., et al. (2013) Side effects of gender- fair language: How feminine job titles influence the evaluation of female applicants. *European Journal of Social Psychology* Vol. 43, No. 1, pp. 62-71.

16: Horvath, L., & Sczesny, S. (2015) Reducing women's lack of fit with leadership positions? Effects of the wording of job advertisements. *European Journal of Work and Organisational Psychology* Vol. 25, No. 2, pp. 316-328.

17: Verweken, D., et al. (2015) Warm-hearted businessmen, competitive housewives? Effects of gender-fair language on adolescents' perceptions of occupations. *Frontiers in Psychology*.

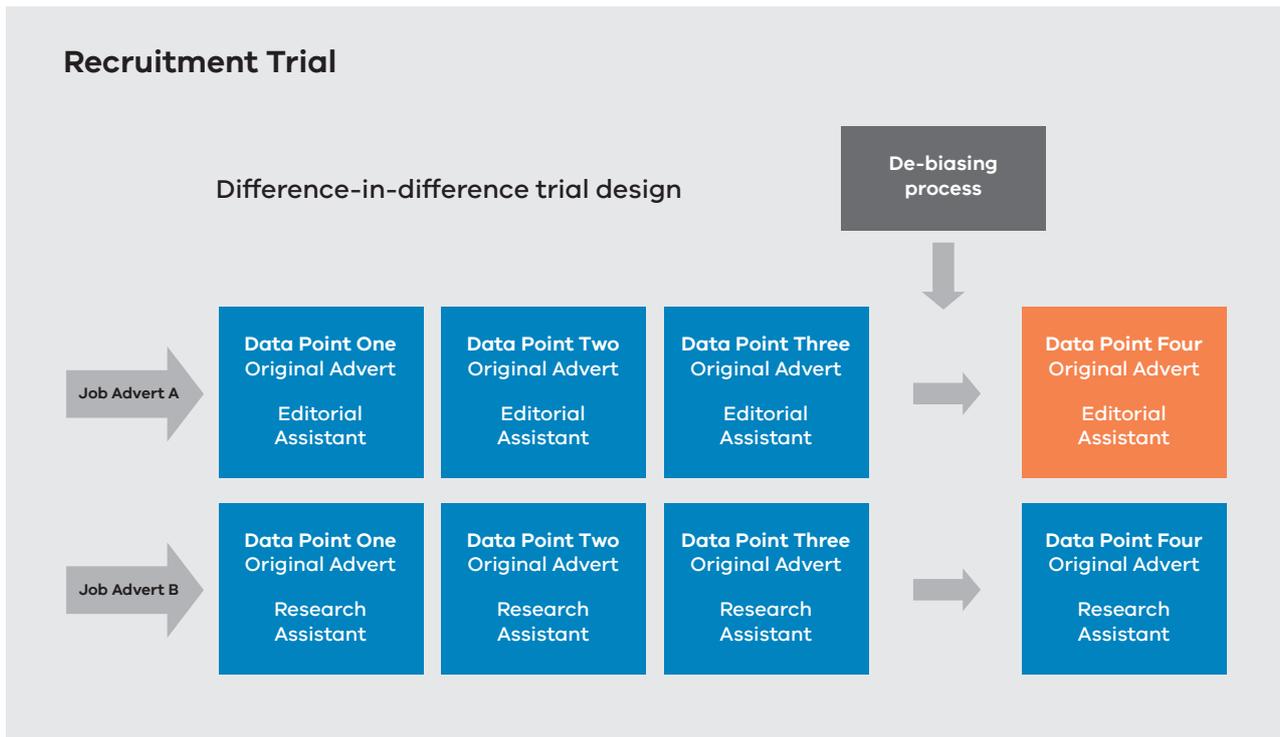


Figure 4

So, generally, if more women applied to job openings in May than in November, this effect should be apparent across all advertisements, not just the debiased ones.

This study design meant all applicants would experience a standard application experience: they would see the job advertised, they will apply for the job and they will go through a standard requirement process once they have applied for the role.

A matched job advertisement would have four components:

- salary (similar amount of money)
- seniority (e.g., both must be at the same level—both managers or both assistants)
- similar field (e.g., both must be in engineering)
- similar proportions of previous applicants (the number of applicants that have previously applied for each role would be similar).

The researchers followed a protocol¹⁸ in the process of de-gendering the advertisements:

1. Using an algorithm, they first measured how gender biased the ad was based on the words identified in Gaucher et al, and then aimed to create a gender balanced ad that includes approximately the same number of male-coded and female-coded words.
2. To move from an unbalanced to a balanced advertisement, unnecessary gendered words were eliminated. For example: "You will work collaboratively with our facilities teams" became "You will work with our facilities teams." "Collaborative" is a female-typed word.
3. When a word can't be removed because it needs to be there, a synonym tool was used to find an alternative that is not coded masculine or feminine. The overall goal was to get as close to parity with masculine- and feminine-gendered words as possible.

¹⁸: The researchers used Applied, a software platform designed to address unconscious bias in recruitment, to assist them with generating the masculine and feminine words.

Key points to de-gender language:

- Use a minimisation (remove/replace) principle, trying to avoid “adding” gendered words to equalise. Determine which of the two lists (masculine- or feminine-gendered) is longer and attempt to get the longer list down to the size of the shorter list.
- Focus on the words that are most clearly related to the role and responsibility of the job rather than company descriptions. For example, “(name of organisation) is a market leader and provides thought leadership.” While the general environment matters, the local environment e.g. the role, is more likely to matter to an applicant and defines who the company is looking for.
- In some cases, it’s a general sense of the advert that conveys gender bias. For example, “you will have the opportunity to work with and be one of our leaders within all facets of TS projects” implies that the person needs to be a leader and work with other leaders who are confident, which may deter some women from applying. A suggestion is to replace the sentiment with: “you will have the opportunity to work with colleagues in all facets of TS projects”. While not precisely the same meaning, the treated advert conveys the sense of a more inclusive environment.

Steps to implementation

This research was conducted using the existing systems and processes unique to each partner organisation’s recruitment practices. All organisations that were scoped for the trial had workforces of over 50 employees, and all used recruitment software to manage the recruitment processes, from the initial application through to the interview and final selection of the candidate.

Different organisations developed their recruitment advertisements in different ways. Some ads were developed solely by the human resources departments; some were developed by the team that needed the new employee while others were a combination of both.

Understandably the trial partners were concerned that adding any extra friction to the applicant process, such as extra clicks or added information boxes to gain informed consent, would turn away potential applicants. Balancing the needs of the partners, while meeting the necessary ethical requirements, was critical and impacted on the data collection.

While the simplicity of the trial attracted many candidates during the scoping process, the reality of delivering the trial proved problematic for all organisations. Batches of job advertisements were decoded, and the language was examined, but the organisations were unable to meet all the ethical requirements. This included finding a matched advertisement to post alongside a control ad, while some did not meet the consent requirements (see Box 3). Consequently, no meaningful data was collected. Further, the difference-in-difference trial design meant a large number of adverts was required to draw any conclusions and show any effect,¹⁹ and this became problematic as the organisations did not have the volume required.

Additional barriers for the organisations included:

- an urgent need for tight turnarounds of advertisements. Human resource departments are fast paced, dynamic environments, and many needed to promote job ads as soon as they received them, without waiting for the researchers to de-gender the advert.
- staff changes to key personnel
- the need for clearance and sign off from managers or head office
- clearance from the organisation’s legal department.

¹⁹: While VicHealth’s trial was underway, SEEK, a large online employment market, conducted their own RCT examining the effects of gendered language in job advertisements. The study formed part of their Laws of Attraction, suite of research with a reported sample size of 6000 job applicants and results showing that gendered language has an effect on the candidate pool.

Ethical requirements for potential research subjects (job applicants).

Due to the nature of testing gendered language in the 'real' world, it was not viable to notify the job applicants of the study by the usual methods e.g. a one-page plain language statement describing the research study and a signed consent form. The recruitment specialists who were our trial site partners indicated that this amount of friction will deter candidates from applying for a role.

To ensure information was included and consent was granted, a short statement was included in the position description that advised the organisation was involved in a research project examining recruitment practices.

After the candidate applied for a position, they were sent an email which included the organisation's usual confirmation of application text, plus an additional short statement that allowed them to withdraw from the research if they chose to, by contacting the organisation's Privacy Officer. The Privacy Officer would then remove any candidate's personal information and forward the de-identified data to the researchers for analysis.

To combat the candidates' fear that they would be disadvantaged by seen to be a 'troublemaker' or 'not a team player' for not taking part in the research, the Privacy Officer was determined the best contact point for candidates to withdraw. An accompanying statement stating, "This will not disadvantage your application" was included to reassure candidates.

While these ethical requirements seemed relatively straightforward, and support material was developed to help each partner organisation understand the steps involved in collecting data, many organisations had difficulty meeting every step.

Measuring outcomes

The first set of partners consisted of

- a mining company;
- two Victorian Government statutory authorities;
- a Victorian Government department;
- two large auditing firms.

Each organisation submitted a number of advertisements to be decoded. In all cases except one, the organisations were interested in increasing the number of female candidates to roles that were male-dominated.

There was a high level of gender neutral language used in the majority of the adverts. The organisations that were interested in the trial were not a random sample of all existing possible organisations in Victoria but likely the ones already more committed to gender equality and diversity and thus, more likely than the average company/agency to have introduced decoders or other debiasing tools.

This indicates the use of decoders in the organisations, not unsurprisingly, as the subjects consisted of a large organisations, who could reasonably be expected to use the most current human resource practices.

One organisation submitted 10 job advertisements, half of which had gender neutral language and half of which used masculine language, indicating that the ads were developed differently. Scoping conversations with several organisations indicated that while the organisation may be using a decoder, the human resource manager was not sure how widely it was used across the organisation, or how effective it was.

After the decoding process was completed, the adverts were sent back to the organisations to post. The modifications the trial design underwent to meet ethical requirements, resulted in additional steps needed by the organisations to collect data. This, coupled with the large number of advertisements needed to show effect with the difference-in-difference trial model (see Figure 4 earlier), resulted in a lack of data and no findings.

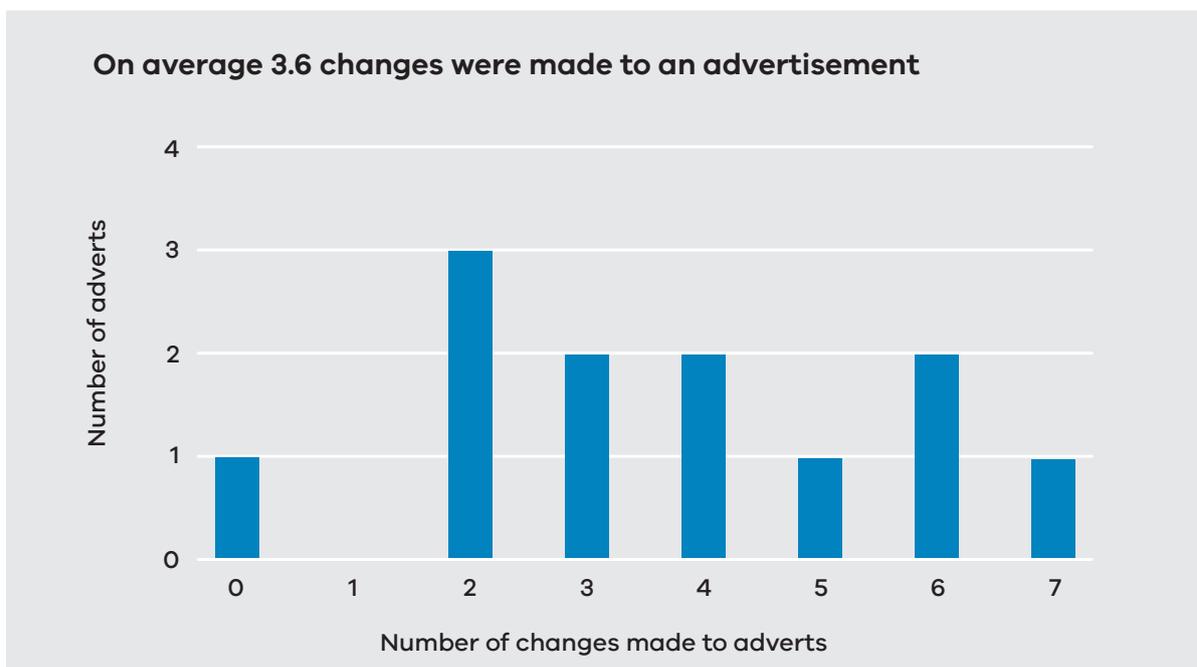


Figure 5

Conclusion

While bigger firms use gender decoders and other software platforms that seek to mitigate the effects of unconscious bias in recruitment, can this knowledge be made accessible to the small- to medium-enterprise sector? These businesses make up 98% of Victoria’s business sector and have fewer resources at their disposal – be it time, information, money or staff. Working with this sector and encouraging gender equitable practices could help to build gender equality in this sector.

Recent work undertaken by VicHealth²⁰ and others in researching healthy masculinities shows that men can benefit from stepping outside rigid gender stereotypes, including stereotypes about the work they ‘should’ do, and how they engage in the workforce. Exploring behavioural insight trials, encouraging men to work part-time or flexibly could impact on increasing numbers of men taking up more caring responsibilities outside of the workplace. Evidence suggests that dismantling rigid gender stereotypes and increasing inclusion and participation will benefit girls and women at all stages of their lives, as well as create a society where men’s and boys’ choices are not limited by gendered social norms.

Occupational segregation also reflects broader social trends. The increasingly common trend of ‘overwork’ perpetuates gender segregation in a number of occupations.²¹ While working long hours is an expected norm in many male-dominated occupations, women, especially mothers, may be less able to meet this expectation as their time is subject to more family demands than their male counterparts, which would affect their job preferences and aspirations.²²

Building gender equality and inclusion in the workforce not only supports and benefits all members of that workforce – female and men - but also makes good business sense. Organisations across a variety of sectors want to benefit from 100% of the talent pool and not deprive themselves of a large number of potential employees, for example, a talented female engineer, purely because they use a heavily ‘masculine’ word in their job ad. The research by Gaucher et al. 2011 provides the evidence for a link between the gendered nature of the language used in job advertisements and the assumptions people make about the job and their likely fit.

20 : Hellman, B., Barker, G. < Harrison, A. (2017) *The Man Box: A study on being a young man in the US, UK and Mexico.*

Retrieved from: <https://promundoglobal.org/wp-content/uploads/2017/03/TheManBox-KeyFindings-EN-Final-29.03.2017-POSTPRINT.v2.pdf>

21: Goldin, C 2014, ‘A grand gender convergence: Its last chapter’, *The American Economic Review*, vol. 104, no. 4, pp. 1091–119.

22: Cha, Y 2013, ‘Overwork and the persistence of gender segregation in occupations’, *Gender & Society*, vol. 27, no. 2, pp. 158–84.

Intervention

3

CV

De-Identification

Unconscious stereotypes relating to personal characteristics, such as gender, age and ethnicity have been shown to be activated by data on applicants' CVs and to bias recruitment decisions.

One approach to reducing the potential bias due to stereotypes is to remove personal data that are unrelated to an applicant's capacity for the role. Although there are risks and barriers to implementing a CV de-identification strategy, an analysis of the effects of CV de-identification outcomes can provide useful insights into unconscious bias in recruitment decisions and may provide valuable insight to help inform other strategies for improving merit-based recruitment and selection processes.

The need For CV de-identification

Recent studies have shown that ancillary details on résumés, such as extra-curricular activities, can signal class and socio-economic status, thus eliciting unconscious bias in recruiters². In a study of hiring practices of top US law firms, applicants who included extra-curricular activities stereotypical of upper class men were overwhelmingly favoured above all others, controlling for other factors. This effect was not found for female applicants.

Researchers have used mock CVs to examine discrimination against certain social groups (e.g. women, ethnically diverse, foreign and older workers) by changing the names of applicants but keeping all other details of the résumés matched. Some studies show effects for race and others for gender. These studies found that:

- Women and people of minority races are less likely to be contacted, despite having the same qualifications as those with Caucasian and male names. This is consistent across the US, Canada, Sweden and Australia, among other countries (Carlsson & Rooth, 2007; Oreopoulos, 2009; Pager et al., 2009).
- For low skilled jobs (e.g. waitress, customer service), Anglo-Saxon candidates received the highest call-back rates (35%) compared to four other ethnic groups: Indigenous Australian, Chinese, Italian, and Middle Eastern. Of the latter, Chinese and Middle Eastern candidates received the lowest call back rates, 21% and 22% respectively (Booth et al., 2012).

De-identification of résumés is intended to be a method of 'levelling the playing field'. By removing identifying characteristics, this strategy aims to remove potential sources of bias for people making hiring decisions.

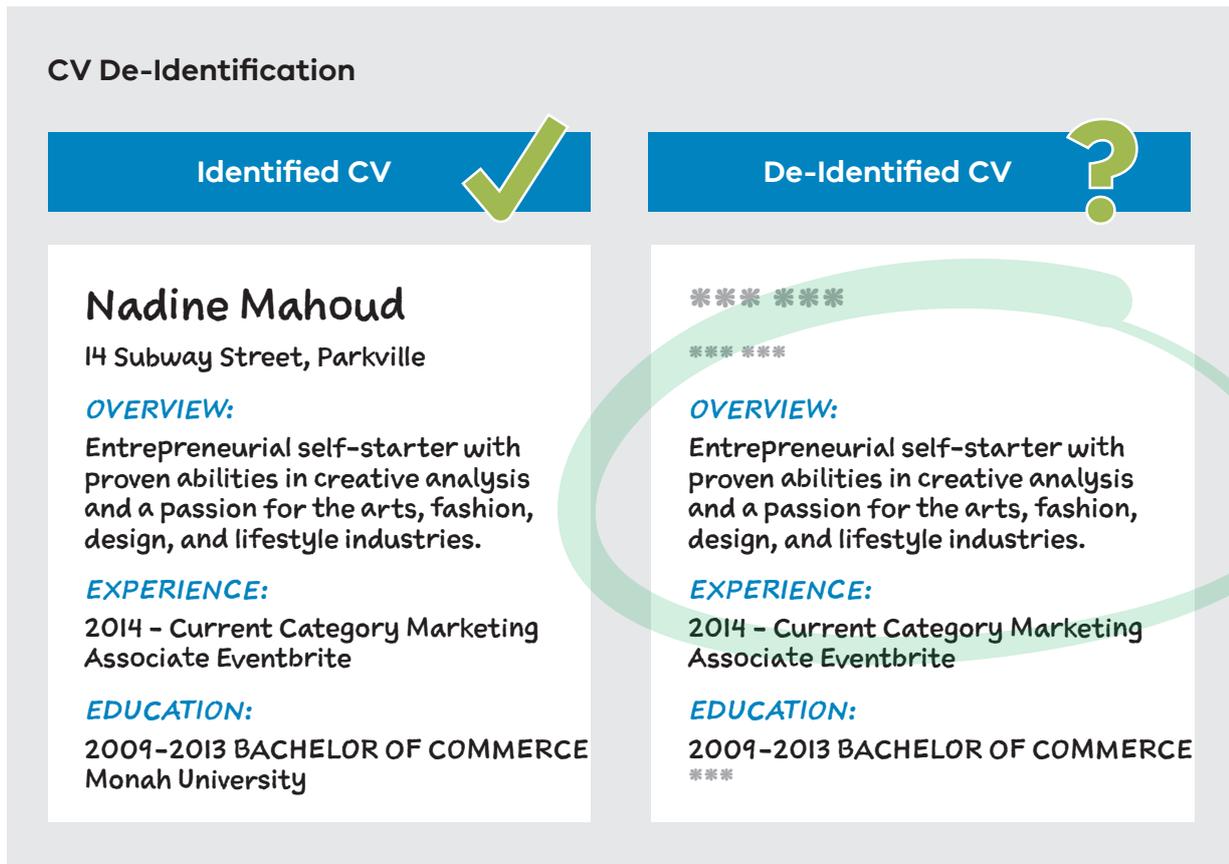


Figure 6. An example of a fictitious curriculum vitae with identified and de-identified socio demographic information. The example on the right only includes merit-based information that is less likely to result in prejudiced attitudes towards the applicant.

What are the risks/barriers?

A common criticism of diversity programs is that they may unfairly disadvantage the majority group. However, the aim of CV de-identification is not to prioritise unsuitable candidates over more qualified ones. Rather, the process is designed to allow consideration for a greater diversity of suitable applicants.

Some companies may want to select from specific social groups (e.g., to increase representation of Aboriginal or Torres Strait Islander peoples). In this case, blind résumés may actually decrease opportunities for people in minority groups.

Expected outcomes

At the interview stage, studies have found CV de-identification benefits both women and ethnically diverse candidates. However, at the job offer stage, this benefit persists largely for women only, and less so for those from minority ethnic backgrounds (Bøg & Kranendonk, 2011; Edin & Lagerström, 2006).

How to measure outcomes

CVs are first evaluated at the shortlisting stage. If more CVs from minority group members are shortlisted under conditions of de-identification compared to when that information is visible, this suggests that anonymized CVs reduce the effects of stereotyping and bias against these groups. See Figure 6 for a comparison of an identified and de-identified version of the same CV.

Case Study on CV De-Identification

VicRoads

VicRoads participated in a 2017 randomised control trial of CV de-identification, to test the effect of de-identification on gender and ethnicity of applicants shortlisted for interview.

Participants

The sample for this dataset comprised 303 applicants (of which 213 were unique) who applied for one of 14 advertised roles between January to April 2017. Demographic variables are summarised in Table 2. Due to missing data, the sample size for analyses including country of birth was reduced ($n = 168$).

Study design

The study included measures of recruitment outcomes when CVs are either identified or de-identified. All applicant CVs were viewed by three hiring managers, one of whom was randomly assigned to view a de-identified version.

Procedures and measures

Under standard hiring practice, three hiring managers independently provide a shortlist of candidates with rankings for interview. The number of candidates shortlisted and ranked is not prescribed and may vary between hiring managers and between roles. The three hiring managers then compare their preferred candidates to make a collective decision regarding which candidates to interview.

In the intervention condition, one of the three hiring managers for each position was randomly assigned to receive de-identified CVs for shortlisting. Each position had a different set of hiring managers.¹

The following demographic information was removed from the de-identified applications: name, gender, ethnicity, and country of birth.²

Predictor variables

We conducted evaluations using identified and de-identified versions of the following participant variables: applicant gender, country of birth, ethnicity of the applicant's name, years of experience, and whether the individual was an internal applicant.

Information regarding participant variables was gathered from submitted job applications. Applicants self-selected their gender identity (with female coded as 2 and male as 1) and country of birth (Australian born coded as 2, and overseas-born coded as 1) in the application process.

1: There was one case where a hiring manager was included on the hiring panel for two different positions.

2: The data collection template for VicRoads asked country of birth but not ethnicity. Therefore a person, for example, of Chinese origin born in Australia would be classified as Australian-born. Ethnicity was inferred from applicant names.

Applicant ethnicity was coded based on the names of the applicant. This was coded independently by two researchers. Where their assessments did not match, a third independent researcher also coded this variable. Ethnicity was either coded as Western European/Caucasian names (1) or non-Western European/Caucasian names (0).

Experience was coded by taking an average of the following experience categories that were responded to in years: experience in similar role; experience in pre-construction for infrastructure projects; experience in coordinating, planning scheduling and resource management related to the development and delivery of infrastructure projects; experience in pre-construction for road infrastructure works; and experience in contract administration. Less than 1 year was coded as 1, between 1 and 3 years was coded as 2, between 3 and 5 years was coded as 3, between 5 and 8 years was coded as 4, and more than 8 years was coded as 5.

Finally, applicants identified themselves as either an internal (coded 2) or external (coded 1) applicant for the role.

Moderator variable

The moderator variable was whether CVs were de-identified, or presented in a fully identified format. In the de-identified condition, hiring managers made their decisions to shortlist using the de-identified CV.

Outcome variable

The outcome variable of interest was whether applicants were shortlisted for interview.

Covariates

The number of applicants for each position, and seniority of the role, were included as covariates in the model. Seniority of the role was coded as one of the following levels: 1 for standard roles and 2 for senior roles.

Analyses

Two research questions were analysed using multiple linear regression models. The first model examined the direct effect of the predictor variables on the shortlisting variable. The second model examined whether relationships between predictor variables and the shortlisting decision were moderated by CV de-identification. Both sets of analyses included covariates.

The two research questions were:

1. Which participant characteristics predict whether an applicant is shortlisted, when all available information is identified?
2. Does de-identification of CVs influence the relationship between applicant characteristics and likelihood of shortlisting?

Results

Demographics

Candidate demographics are provided in Table 2. There was some overlap across roles, with 49 candidates applying for more than one position, and one of the hiring managers evaluating candidates for two roles. As such, although there were 303 candidate evaluations in the intervention, only 213 of these represented unique applicants.

Correlations

Pearson correlations between variables revealed positive associations between country of birth and being an internal applicant, as well as negative correlations between years of experience and gender, external applicants, and being born overseas. progressing further through the recruitment process and male gender, as well as attending a university other than Monash University or the University of Melbourne (Table 3).

Table 2

Demographic Data for Unique Applicants

| Demographic | Number | Percentage |
|---|--------|------------|
| Gender | | |
| Male | 180 | 84.51% |
| Female | 33 | 15.49% |
| Identification as Aboriginal or Torres Strait Islander | | |
| Yes | 0 | 0% |
| No | 213 | 100% |
| Country of birth | | |
| Australia | 57 | 26.76% |
| Elsewhere | 109 | 51.17% |
| No response | 47 | 22.07% |
| Disclosure of disability | | |
| Yes | 2 | 0.94% |
| Not applicable | 209 | 98.12% |
| Prefer not to respond | 2 | 0.94% |
| Applicant | | |
| Internal candidate | 56 | 26.29% |
| Internal applicant avg. merit score | 2.95 | |
| External candidate | 157 | 73.71% |
| External applicant avg. merit score | 3.44 | |

N = 213

Research Question 1

Which applicant characteristics predict the progress of an applicant through the recruitment process when all available information is identified?

Research Question 1 was analysed using a logistic regression model, with number of applicants and seniority of the role as covariates.

Results of the regression model demonstrated positive and significant effects of the predictor variables on the odds of being shortlisted for interview. Odds of being shortlisted were 2.24 ($p < .001$) greater for internal applicants compared to external applicants, and 1.8 ($p < .015$) greater for Australian-born applicants than overseas-born applicants. Odds of being shortlisted were also 1.38 ($p < .001$) greater for candidates with more years of experience compared to candidates with less experience. Understandably, the more applicants per job, the lower the odds were of getting shortlisted (Odds ratio = .67, $p < .001$). There was no difference between females and males on odds of being shortlisted ($p = .533$), nor was job seniority associated with shortlisting ($p = .196$).

Research Question 2

Does de-identification of curriculum vitae influence the relationship between applicant characteristics and likelihood of shortlisting?

Research Question 2 was analysed using a multilevel logistic regression model, with number of applicants and seniority of the role as covariates. This model was run twice, once for gender and once for country of birth.

Results of the regression model demonstrated that de-identifying applicant country of birth had a significant effect on shortlisting for interview. Australian-born applicants were preferred when this information was visible, however if this information was de-identified, overseas-born applicants were preferred ($\beta = -1.21, p < .005$). This relationship is visualised in Figure 7.

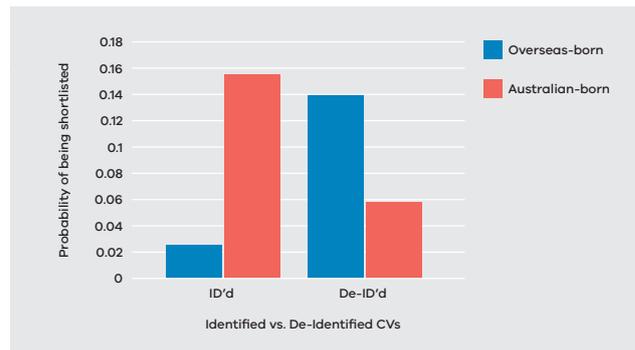


Figure 7. The interaction of applicant country of birth with de-identification, demonstrating that applicants born in Australia enjoy a greater probability of being shortlisted before de-identification, while overseas-born applicants have a greater probability of being shortlisted under de-identified conditions.

Table 3

Pearson correlations between participant gender, internal/external applicant, country of birth, years of experience, experimental condition, shortlisting outcome, and number of applicants for the role.

| Variables | Mean | SD | 1 | 2 | 3 | 4 | 5 | 6 |
|--|------|------|--------|--------|--------|------|-----|--------|
| 1. Participant gender | 1.12 | .33 | | | | | | |
| 2. Internal/External Applicant | 1.26 | .44 | .03 | | | | | |
| 3. Country of birth | 1.29 | .45 | -.02 | .27** | | | | |
| 4. Merit variable | 3.36 | 1.28 | -.12** | -.18** | -.22** | | | |
| 5. Baseline or De-Identified Condition | .33 | .47 | .00 | .00 | .00 | .00 | | |
| 6. Shortlisting outcome | .28 | .45 | -.01 | .27** | .13** | .08* | .01 | |
| 7. Number of applicants | 3.54 | 1.48 | -.01 | -.28** | -.13** | .02 | .00 | -.35** |

Notes. Gender: male = 1, female = 2; 1 = external applicant, 2 = internal applicant; 1 = overseas-born applicant, 2 = Australian-born applicant; 0 = Baseline, 1 = De-identified; 0 = Not shortlisted, 1 = Shortlisted. N = 714-909. * $p < .05$, ** $p < .01$.

Discussion, implications and conclusions

VicRoads was the only organisation to undertake CV de-identification using a randomised control trial approach.

In this evaluation, we found evidence that country of birth, internal vs external applicant, and experience influenced an applicant's likelihood of being shortlisted.

Gender was not a significant predictor, but the small number of female applicants meant that tests to detect the difference had low statistical power. The effect of country of birth on shortlisting was affected by de-identifying CVs. Overseas-born applicants had a higher probability of being shortlisted for interview when country of birth was de-identified, compared to when this information was available to hiring managers.

Only a small number of overseas-born applicants did not have appropriate work rights, ruling out the possibility that this result is due to work visa issues. Results also do not suggest that there is a racial bias at play against overseas-born applicants. Ethnicity of name did not reveal a preference for Caucasian (Western European) names compared to non-Caucasian names.

The result may be driven by perceptions of the applicant's English language skills, and/or having qualifications from an overseas university, consistent with prior research findings (see Oreopoulos, 2009). However, this explanation cannot be tested using the data collected in this intervention. Again, canvassing hiring managers' perceptions of overseas-born applicants may provide insights into the reasons for this effect.

Contrary to our results, prior research has shown that women tend to progress further through recruitment when their gender identity is removed (e.g., Behaghel, Crépon, & Le Barbanchon, 2012). As mentioned, the small number of female applicants for the male dominated project engineering role in the current study may have affected the result.

Without a diverse pool of applicants, de-identification will not have an impact on the diversity of shortlists, highlighting the limitations of focusing diversity efforts at a single stage of recruitment. The lack of potential female applicants contrasts with the supply of overseas-born applicants, for whom de-identification increased the likelihood of shortlisting. Organisations planning to undertake CV de-identification are urged to consider the demographic composition of the current applicant pool, and whether changes in job advertisement language and placement might complement this intervention by increasing diversity in the applicant pool before de-identification is applied.

Case Study on CV De-Identification

Hall & Wilcox

Participants

The sample comprised 1,993 applicants (1148 female, 822 male, 23 unspecified) who applied for a seasonal clerk role at Hall & Wilcox between 2014 and 2017. The baseline condition comprised 1,069 applicants from 2014 and 2015, with 924 applicants from 2016 and 2017 in the de-identified group. Due to a data collection error in 2017, data from this cohort was excluded from analyses, with the modified sample size comprising $N = 1,510$ applicant (910 female, 595 male, 5 unspecified).³

Demographic variables are summarised in Table 4. Complete data for gender were obtained for $n = 1,505$ participants, $n = 1,498$ participants for university attended data, and $n = 1,501$ applicants whose names were able to be coded for ethnicity of origin. As such, some of the analyses contain samples smaller than the overall number of participants in this project.

Study design

The study included measures of recruitment outcomes prior to and following the introduction of CV de-identification. In the de-identified condition, all hiring managers made shortlisting decisions based on the de-identified version of the applicants' CV.

Procedures and measures

Hall & Wilcox is an independent business law firm operating throughout Australia. They conducted a program of CV de-identification from 2014 – 2017, with the goal of increasing gender diversity in their seasonal clerkship program. The intervention involved using an online recruitment tool to de-identify certain demographic information relating to candidates during the hiring process. Baseline

data with standard identification was collected for a seasonal clerk role over 2014 and 2015. Throughout 2016 and 2017, the following information was de-identified in all applications: Applicant name, applicant home address, and any reference to applicant gender (e.g., pronouns). The data was provided by Hall & Wilcox to the Recruit Smarter team as a pre-existing data set in archival form, as it was collected prior to the commencement of the program. The variables collected and used in the present analysis are described below.

Predictor variables

We conducted evaluations on the following two variables of interest: applicant gender and applicant ethnicity.

Applicant gender were inferred from candidate name, as coded independently by two researchers. Where their assessments did not match, a third independent researcher also coded these variables.

Applicant ethnicity was coded based on the names of the applicant. As with gender, this was coded independently by two researchers. Where their assessments did not match, a third independent researcher also coded these variables. A coding scheme was devised to code for the perceived ethnicity, divided into several regional areas. Final coding categories included 1 = Western/Caucasian names (Western and Northern European), 2 = Western Caucasian names (Eastern and Southern European), 3 = Caucasian first name with non-Western surname, 4 = East and South-East Asian names, and 5 = Other non-Caucasian names (including Middle Eastern, African, and Indigenous Australian names). This coding scheme was developed using both existing experimental data on racial discrimination in hiring in an Australian context (Booth et al., 2012), and incorporating historical patterns of migration to Victoria.

³: Data analysis was run with and without the 2017 data included; interpretation of results did not vary between the two sets of analyses, except as noted.

Moderator variable

The moderator variable was whether the CVs were de-identified or remained in identified form. CV de-identification was integrated into the existing Hall & Wilcox hiring process. In the standard hiring practice, applicants who submit CVs for a role are reviewed by hiring managers, who shortlist a number of candidates for interview. After the interview successful applicants are offered roles. De-identification of CVs occurred between the application and shortlisting for interview stage.

Covariate variable

Applicant higher education institution was not de-identified, however was included in the model as a covariate. Higher education institution attended was coded as a binary variable based on whether the candidate attended either The University of Melbourne or Monash University, versus another university. Across government roles, alumni of the University of Melbourne and Monash University appear to enjoy preferential hiring outcomes.

Outcome variable

The outcome of interest was how far candidates proceeded through the recruitment process (Not Shortlisted, Shortlisted, Interviewed, and Hired).

Table 4

Gender, University Attended, and Ethnicity, presented according to Recruitment Year

| | 2014 | 2015 | 2016 |
|---|-------------|-------------|-------------|
| Gender | | | |
| Female | 309 (59.8%) | 333 (60.3%) | 268 (60.8%) |
| Male | 206 (39.8%) | 217 (39.3%) | 172 (39.0%) |
| Unspecified | 2 (0.4%) | 2 (0.4%) | 1 (0.2%) |
| University | | | |
| Monash/Melbourne | 289 (55.9%) | 284 (51.4%) | 233 (52.8%) |
| Deakin, LaTrobe & other | 225 (43.5%) | 260 (47.1%) | 207 (46.9%) |
| Missing | 3 (0.6%) | 8 (1.4%) | 1 (0.2%) |
| Ethnicity of name | | | |
| Names of Western and Northern European origin | 369 (71.4%) | 362 (65.6%) | 287 (65.1%) |
| Names of Central and Eastern European origin | 55 (10.6%) | 68 (12.3%) | 50 (11.3%) |
| Western origin first name and East Asian origin last name | 45 (8.7%) | 55 (10.0%) | 55 (12.5%) |
| Names of East Asian origin | 15 (2.9%) | 21 (3.8%) | 14 (3.2%) |
| Names of other origin (Middle Eastern, South Asian, African, Indigenous Australian) | 29 (5.6%) | 43 (7.8%) | 33 (7.5%) |
| Missing | 4 (0.8%) | 3 (0.5%) | 2 (0.5%) |

4: The data was also analysed using a set of logistic regression models, with binary outcomes for each of the selection stages (reached vs. not reached). Results corresponded to those of the linear regression model presented in the current report. Linear regression is presented as it allowed for analyses of recruitment outcomes in a single model.

Analyses

Each research questions was analysed using a linear regression model.⁴ The model examined the direct effect of the predictor variable (gender or perceived ethnicity) on participant progress through the recruitment process. The model also examined whether any direct relationships between predictor variables and the recruitment outcome varied depending on whether CVs were de-identified. The model was run twice for each predictor, initially without covariates, and a second time including covariates.

The two research questions were:

1. Does de-identification of gender equalise recruitment outcomes between men and women?
2. Does de-identifying ethnicity information by removing applicant full names from CVs equalise recruitment outcomes for applicants from ethnically diverse backgrounds?

Results

Demographics

Table 4 displays participants' demographic data. Overall, there was a slightly larger proportion of females than male applicants, particularly in the baseline condition. Approximately half of the applicants attended Melbourne or Monash university, with the majority of individuals having a Western or Caucasian sounding name.

Correlations

Pearson correlations between variables revealed positive associations between progressing further through the recruitment process and male gender, as well as attending Monash University or the University of Melbourne (Table 5).

Table 5

Pearson correlations between participant gender, ethnicity, and university attended (Melbourne and Monash vs other), experimental condition, and application stage reached.

| Variables | Mean | SD | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|--|------|------|-------|------|-------|-------|---|---|---|
| 1. Participant gender | .40 | .49 | | | | | | | |
| 2. Ethnicity | 1.70 | 1.21 | -.04 | | | | | | |
| 3. University | .54 | .50 | .03 | .01 | | | | | |
| 4. Baseline or De-Identified Condition | .29 | .45 | .04* | .02 | -.002 | | | | |
| 5. Selection Outcome | .35 | .81 | .06** | -.04 | 1.4** | .32** | | | |

Notes. Gender: 0 = female, 1 = male; 1 = Western/Caucasian names (Western and Northern European), 2 = Western Caucasian names (Eastern and Southern European), 3 = Caucasian first name with non-Western surname, 4 = East and South-East Asian names, and 5 = Other non-Caucasian names 0 = Non-Melb or Monash Uni, 1 = Melb or Monash Uni; 0 = Baseline, 1 = De-identified; N = 1,956 – 1,993. * p < .05, ** p < .01.

Research Question 1

Does de-identification of gender equalise recruitment outcomes between men and women during recruitment?

Research Question 1 was analysed using a moderated linear regression model. The model examined the direct effect of gender on recruitment outcome, and whether this relationship was moderated by the intervention (i.e., baseline or de-identification condition). The model was also run a second time to include the additional two predictor variables (applicant ethnicity and university attended) as covariates.

The overall linear regression model was significant, $F(3, 1501) = 3.31, p = .020$. However, results indicate no main effect of gender on recruitment outcomes, $t(1501) = .52, p = .602$, and no interaction between gender and de-identification to predict recruitment outcomes, $t(1501) = .85, p = .393$. There was little difference in the model when covariates were included.

Research Question 2

Does de-identification of applicant's name on submitted curriculum vitae improve probability of progressing through recruitment stages for applicants from culturally diverse backgrounds?

Research Question 2 was analysed using a moderated linear regression model. The model examined the direct effect of candidate ethnicity on progress through the recruitment process, and whether this relationship was moderated by the intervention (i.e., baseline or de-identification condition). The model was also run a second time to include covariates. These included participant gender and university attended.

The overall linear regression model was significant, $F(3, 1497) = 5.11, p = .002$. Results indicated a main effect of candidate ethnicity on recruitment outcome, $t(1497) = -2.12, p = .034$. Candidates with Caucasian and European-origin names progressed further through the recruitment process than candidates with names perceived as non-European and culturally diverse. The one exception was candidates with East Asian names (first and last), who tended to progress furthest in the recruitment process. De-identification did not significantly moderate the relationship between candidate ethnicity and recruitment outcome $t(1497) = -.17, p = .868$, suggesting that de-identifying data did not reduce the tendency for individuals with Caucasian names to proceed further through the recruitment process. Inclusion of covariates resulted in minimal change to the model, however this analysis indicated that applicants from Melbourne and Monash universities tended to proceed substantially further through the recruitment process compared to applicants from other universities, $t(1497) = 5.80, p < .001$.

Discussion

The intervention conducted at Hall & Wilcox tested whether CV de-identification affected how far diverse candidates progressed through the hiring process. The variables de-identified in applicant CVs included candidate name and gender. We found that there was no significant difference in recruitment outcomes for men and women. However, candidates with names that implied Caucasian cultural heritage progressed further through the recruitment process than applicants with names that indicated a non-Caucasian cultural background. For both name and gender, subsequent CV de-identification did not significantly change the pattern of results.

Minimal gender differences were noted in the progression of men and women through recruitment, both at the baseline and de-identified condition. This result may be due to the non-gendered nature of the seasonal clerk role, or the relatively similar numbers of men and women applying for roles. In addition, gender disparity in hiring tends to occur in higher-ranking roles, such as management and executive jobs, roles which were not captured in this pilot program. Results may have also been complicated by dissimilar numbers of women applying across years; for example, the sample comprised of over 60% female applicants in 2016 but dropped back to just over 50% in 2017.

One finding of note was that while women typically comprised the majority of initial applicants across all years of the trial, this majority was not reflected in subsequent stages of recruitment (i.e., more women than men applied for the role, but equal numbers were hired; see Appendix A). So, while there does not appear to be bias towards either gender from shortlisting onwards, there is a greater drop-out rate of female candidates compared to male candidates between the application and shortlisting stage. As such, Hall & Wilcox may want to consider whether more support could be given to female applicants to improve their likelihood of shortlisting, which would better reflect the gender composition of the underlying applicant pool. For example, female applicants may not be as confident in expressing their skills and expertise or feel immodest in talking up their achievements.

These types of gender differences in communication could be noted in the job application process, to encourage female applicants to consider using stronger language in describing their achievements.

The perceived ethnicity of applicants' names was related to progression through recruitment, with candidates whose names implied Caucasian cultural heritage more likely to progress further through the recruitment process than individuals with less 'Western-sounding' names. The one exception was applicants with an East Asian first and last name, who performed equal or better than applicants with Western-origin names. However, this result is based on a very small sample of applicants with East-Asian names, and thus should be viewed as preliminary only.

The relationship between ethnicity of names and recruitment outcomes was not substantially changed by the introduction of CV de-identification. There are two possible explanations for this finding. First, this may reflect other difference between applicants that are related to cultural knowledge and capital. For example, applicants from non-Western backgrounds for whom English is a second language may have less understanding about how to answer qualitative questions about their skills, abilities, and values in ways that increase their likelihood of progressing through the application process. Applicants from non-Western backgrounds may have responded to questions in a slightly different manner compared to applicants from a Western background or may have emphasised values that are more highly prized in non-Western cultures compared to those valued in a Western cultural context. This may place them at a disadvantage when applying for roles at Australian law firms.

Second, there may be a merit gap between Western and non-Western applicants, potentially on the basis of poorer English-language skills, and perhaps relatedly, university results. This information was not available in the present investigation, however Hall & Wilcox may test for evidence of this hypothesis in subsequent years of recruitment by collecting objective information about applicant academic abilities, such as GPA scores.

Additionally, software capable of calculating English-language skills may provide evidence for or against the hypothesis that English expression is contributing to the poorer recruitment outcomes of applicants with non-Western names. Understanding the mechanism of this effect is important in assisting Hall & Wilcox and other organisations in determining how best to support applicants from underrepresented backgrounds in their job applications.

Case Study on CV De-Identification

Department of Premier and Cabinet

Participants

The sample for this dataset comprised 796 applicants (378 female, 390 male, 28 unspecified) who applied for one of 30 roles within the Department of Premier and Cabinet (DPC) in May and October 2017 and from May through June 2018. The baseline condition comprised 499 applicants, with 297 applicants in the de-identified group. Demographic variables are summarised in Table 5. Complete data for gender was obtained for $n = 768$ participants, with $n = 696$ participants providing suburb data and $n = 785$ participants providing university attended data. As such, some of the analyses contain samples smaller than the overall number of participants in this project.

Study design

The study included measures of recruitment outcomes prior to and following the introduction of CV de-identification. In the de-identified condition, all hiring managers made shortlisting decisions based on the de-identified version of the applicants' CV.

Procedures and measures

The intervention used an online recruitment tool to de-identify certain demographic information relating to candidates during the hiring process. Baseline data with standard identification was collected in May 2017 ($n = 341$) and May 2018 ($n = 158$). Throughout October 2017 ($n = 246$) and June 2018 ($n = 51$), the following information was de-identified in all applications: applicant name, applicant home address, applicant university/institution, applicant degree type, and referee name and organisation. In May and June 2018, the Department of Premier and Cabinet also collected information about the applicant's self-identified ethnic/cultural heritage, their parents' place of birth, and languages spoken other than English (all de-identified to hiring managers).

Predictor variables

We conducted evaluations on the following three variables of interest: applicant gender, applicant home address, and applicant university attended. The latter two variables were taken to reflect social capital.

Applicants were coded as male or female based on their self-selected title (e.g., Mr, Ms). Where applicants used a non-gendered title, such as Dr, gender was coded as missing data and excluded from gender analyses. Home address (suburb only) was coded from 1 to 10 using the Australia Bureau of Statistics' (2018) Index of Relative Socio-Economic Advantage and Disadvantage (IRSAD). A low score (minimum = 1) indicates relatively greater disadvantage and a lack of advantage in general, while a high

score (maximum = 10) indicates a relative lack of disadvantage and greater advantage in general. Higher Education institution attended was coded as a binary variable based on whether the institution attended by the candidate was a member of the Group of Eight⁵ (Go8; considered a marker of prestige), or otherwise. Across government roles, alumni of Go8 institutions – especially the University of Melbourne and Monash University – appear to enjoy preferential hiring outcomes. We thus wanted to investigate whether de-identifying this information influenced recruitment progress for applicants from other tertiary institutions.

Moderator variable

CV de-identification was integrated into the existing DPC hiring process. In DPC's standard hiring practice, applicants who submit CVs for a role are reviewed by hiring managers who shortlist a number of candidates for interview. After the interview, referee checks are conducted and successful applicants are offered roles. De-identification was in place for CVs between the application and shortlisting for interview stage.

Outcome variable

The outcome of interest was how far candidates proceeded through the recruitment process. This was measured on a four-point scale that included

- 1 = Application;
- 2 = Interview;
- 3 = Talent Pool;
- 4 = Job Offer.

Analyses

Three research questions were analysed using moderated linear regression model. The model examined the direct effect of the predictor variable (gender, suburb, or university attended) on the recruitment outcome variable. The model also examined whether any direct relationships between predictor variables and the recruitment outcome were moderated by CV de-identification. The model was run twice for each predictor, initially without covariates, and a second time including covariates.

The three research questions were:

1. Does de-identification of gender reduce the hiring gap between men and women during recruitment?
2. Does removing home addresses from CVs equalise recruitment outcomes for applicants from socio-economically diverse suburbs?

3. Does de-identification of higher education institution improve recruitment equality for applicants from institutions of varying prestige?

Results

Demographics

Table 6 displays participant demographic data. Overall, there was an even number of males and females, with a slightly higher proportion of males in the baseline compared to the de-identified condition. Suburb relative advantage and disadvantage was similar for participants in the baseline and de-identified groups, while there was a larger proportion of participants from Go8 universities in the de-identified condition than in the baseline condition.

Correlations

Pearson correlations between variables revealed significant negative correlations between male gender and: Living in a more socio-economically advantaged suburb, attending a Go8 university, and being in the de-identified experimental condition. Meanwhile, there were positive correlations between attending a Go8 university and living in a more socio-economically advantaged suburb, and between attending a Go8 university and membership of the de-identified condition (see Table 7.)

Research Question 1

Does de-identification of gender reduce the hiring gap between men and women during recruitment?

Research Question 1 was analysed using a moderated linear regression model. The model examined the direct effect of gender on recruitment outcome, and whether this relationship was moderated by the intervention (i.e., baseline versus de-identified condition). The model was also run a second time to include the additional two predictor variables (applicant home address and university attended) as covariates.

The overall linear regression model was not significant, $F(3, 755) = .75, p = .521$. Results indicate no main effect of gender on recruitment outcome, $t(755) = -1.24, p = .217$, and no interaction between gender and de-identification to predict recruitment outcomes, $t(755) = .14, p = .887$. There was little difference in the model when covariates were included.

5: Go8 universities include The University of Adelaide, The Australian National University, The University of Melbourne, Monash University, The University of New South Wales, The University of Queensland, The University of Sydney, and The University of Western Australia

Research Question 2

Does removing home addresses from CVs equalise recruitment outcomes for applicants from socio-economically diverse suburbs?

Research Question 2 was analysed using a moderated linear regression model. The model examined the direct effect of candidate suburb on recruitment outcome, and whether this relationship was moderated by the intervention (i.e., baseline versus de-identified condition). The model was also run a second time to include covariates. These included participant gender, and whether the participant attended a Go8 university.

The overall linear regression model was significant, $F(3, 683) = 2.90, p = .034$. Results indicated a main effect of candidate suburb on recruitment outcome, $t(683) = 2.92, p = .004$. Visual inspection of the plotted data showed that when participant suburb is not de-identified, participants from higher socio-economic suburbs proceeded further through the recruitment process. De-identification marginally moderated the relationship between suburb and recruitment outcome, $t(683) = -1.81, p = .068$, suggesting that de-identifying data may have reduced the tendency for individuals from higher socio-economic suburbs to proceed further through the recruitment process. When covariates were included in the model, the moderating effect of de-identification on suburb met the threshold for statistical significance, $t(658) = -2.01, p = .045$. This finding was driven by a significant positive relationship between applicant suburb and recruitment progress in the baseline but not the experimental condition (see Figure 8).

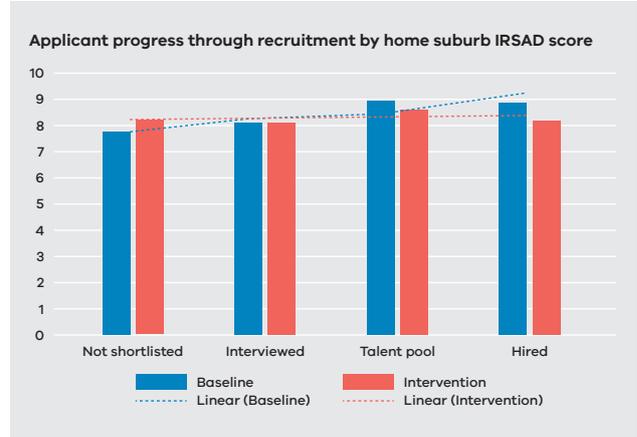


Figure 8. The average IRSAD score of the pool of applicants who reach each stage of the recruitment process, separated by the condition (baseline or intervention). Linear trendlines have been included to demonstrate the increasing socioeconomic status of applicants in the baseline condition.

Table 6
Participant demographic data for the overall DPC sample

| | | Baseline | De-identified |
|-------------------|-------------------------------|-------------|---------------|
| | Total participants | 420 | 304 |
| | <i>Female</i> | 229 | 127 |
| | <i>Male</i> | 184 | 165 |
| | <i>Unspecified Gender</i> | 7 | 12 |
| Suburb | <i>Average IRSAD score</i> | 8.01 (2.43) | 8.22 (2.52) |
| | <i>No. people no response</i> | 43 (10.2%) | 39 (12.8%) |
| University | <i>Go8</i> | 129 (29.7%) | 164 (53.9%) |
| | <i>Non-Go8</i> | 293 (70.3%) | 135 (44.4%) |
| | <i>Missing</i> | 3 | 5 |

Research Question 3

Does de-identification of higher education institution improve recruitment equality for applicants from institutions of varying prestige?

Research Question 3 was analysed using a moderated linear regression model. The model examined the direct effect of candidate university attended (Go8 versus other), and whether this relationship was moderated by the intervention (i.e., baseline or de-identification condition). The model was also run a second time to include covariates. These included participant gender, and participant home suburb socio-economic status.

The overall linear regression model was not significant, $F(3, 771) = 2.31, p = .075$. However, results indicated a main effect of attending a Go8 university on recruitment outcome, $t(771) = 2.55, p = .011$, with visual inspection of the plotted data showing that candidates from Go8 universities progressed further through the recruitment process than their non-Go8 counterparts, when participant information was not de-identified (see Figure 9a-b). De-identification did not significantly moderate the relationship between university and recruitment outcome, $t(771) = -1.16, p = .245$, suggesting that de-identifying data did not reduce the tendency for individuals from Go8 universities to proceed further through the recruitment process. When covariates were included in the model, the main effect of Go8 university was no longer significant, $t(658) = 1.94, p = .053$.

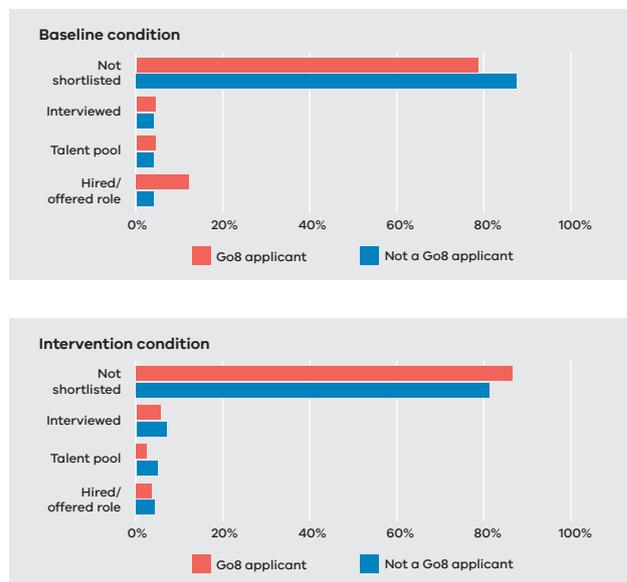


Figure 9a and 9b. The percentage of applicants from Group of 8 or non-Group of 8 universities who reach each stage of the application process relative to the number of applicants in total. Baseline data is shown in Figure 9a, while intervention data is shown in 9b.

Table 7

Pearson correlations between participant gender, suburb IRSAD score, Go8 university attended or not, experimental condition, and application stage reached.

| Variables | Mean | SD | 1 | 2 | 3 | 4 | 5 |
|--|------|------|--------|-------|-------|-------|---|
| 1. Participant Gender | .51 | .50 | | | | | |
| 2. Suburb | 8.14 | 2.41 | -.12** | | | | |
| 3. Go8 University | .40 | .49 | -.09** | .19** | | | |
| 4. Baseline or De-Identified Condition | .37 | .48 | -.09** | .04 | .25** | | |
| 5. Application Stage Reached | .31 | .79 | -.05 | .09* | .08* | -.005 | |

Notes. Gender: female = 0, male = 1; 0 = Non-Go8, 1 = Go8; 0 = Baseline, 1 = De-identified; Application stages include 0 = Not shortlisted, 1 = Interview, 2 = Talent Pool, and 3 = Hired. N = 608-716. * $p < .05$, ** $p < .01$.

Discussion

We tested whether CV de-identification affected how far diverse candidates progressed through the recruitment process. The variables de-identified included candidate gender, home suburb socioeconomic status, and university attended. We found that after de-identification, more applicants from lower socio-economic suburbs progressed further through the recruitment process. Similar results were found for candidates from less prestigious universities, however the effect of de-identification did not reach the threshold of significance. De-identifying did not lead to changes in gender composition of applicants at any stage of the hiring process.

Applicant home address may be an indicator of socioeconomic status, or otherwise indicate an applicant's level of social capital. When this information was identified, there was a relationship between the relative social advantage of applicant home suburbs and the stage of recruitment they reached. In contrast, when this information was removed, the relationship between applicant suburb and stage of recruitment was eliminated. This suggests that hiring managers may be using an applicant's home address to form a prejudicial belief (at a conscious or unconscious level) about what sort of person the applicant is. However, it is also possible that there are additional confounding variables that may explain this relationship, including education level. Furthermore, across the entire dataset there was a higher likelihood of applicants from high socioeconomic suburbs applying for roles at the Department of Premier and Cabinet. The organisation may want to consider different means of attracting applicants from a greater diversity of socioeconomic backgrounds to ensure greater representation of the Victorian population.

Applicants who attended more prestigious universities tended to progress further through the recruitment process compared to those who attended less prestigious institutions. While de-identifying this data trended towards a reduction in the advantage enjoyed by applicants with more prestigious education, this effect did not reach significance.

De-identifying an applicant's gender did not significantly impact the progress of men and women through the recruitment process. This result may be due to the non-gendered nature of the roles advertised, or the relatively similar numbers of men and women applying for roles. In addition, gender disparity in hiring tends to occur in higher-ranking roles, such as management and executive jobs, roles which were not captured in this pilot program. We may expect to see greater benefits to suitably-qualified women applying for higher ranking roles when CVs are de-identified.

Overall, CV de-identification appears to be of particular importance when it comes to indicators of social capital and advantage. Job applicants from more prestigious universities or who live in suburbs with higher socioeconomic status typically progress further in the recruitment process. When this information is removed, applicants from lower socioeconomic suburbs tend to benefit and perform better in recruitment. However, there is no difference to the gender composition of applicants at all hiring stages after de-identification is applied.

Case Study on CV De-Identification

Department of Treasury and Finance

Participants

The sample for this dataset comprised 162 applicants (44 female, 67 male, 51 unspecified) who applied for one of 7 roles within the Department of Treasury and Finance (DTF) from October 2017 through to January 2018. All 7 roles were traditionally male dominated. The baseline condition comprised 42 applicants across 2 roles, with 120 applicants across 5 roles in the de-identified group.

Demographic variables are summarised in Table 5. Complete data for gender was obtained for $n = 111$ participants, and $n = 158$ participants provided university attended data. As such, some of the analyses contain samples smaller than the overall number of participants in this project.

Study design

The study included measures of recruitment outcomes prior to and following the introduction of CV de-identification. In the de-identified condition, all hiring managers made shortlisting decisions based on the de-identified version of the applicants' CV.

Procedures and measures

The intervention involved using an online recruitment tool to de-identify certain demographic information relating to candidates during the hiring process. Baseline data with standard identification was collected in October 2017. From December 2017, the following information was de-identified in all applications: applicant name, applicant gender, applicant university/institution, and the applicant's citizenship status.

Predictor variables

We conducted evaluations on the following two variables of interest: applicant gender and applicant university attended.

Applicants were coded as male or female based on their self-selected title (e.g., Mr, Ms). Where applicants used a non-gendered title, such as Dr, gender was coded as missing data and excluded from gender analyses. Higher education institution attended was coded as a binary variable based on whether the institution attended by the candidate was a member of the Group of Eight⁶ (Go8; considered a marker of prestige), or otherwise. Across government roles, alumni of Go8 institutions – especially the University of Melbourne and Monash University – appear to enjoy preferential hiring outcomes. We thus wanted to investigate whether de-identifying this information influenced recruitment progress for applicants from other tertiary institutions.

6: Go8 universities include The University of Adelaide, The Australian National University, The University of Melbourne, Monash University, The University of New South Wales, The University of Queensland, The University of Sydney, and The University of Western Australia

Moderator variable

CV de-identification was integrated into the existing DTF hiring process. In DTF’s standard hiring practice, applicants who submit CVs for a role are reviewed by hiring managers who shortlist a number of candidates for interview. After the interview, referee checks are conducted and successful applicants are offered roles. De-identification was in place for CVs between the application and shortlisting for interview stage.

Outcome variable

The outcome of interest was how far candidates proceeded through the recruitment process. This was measured on a three-point scale that included Application Stage, Short-listed for Interview, and Job Offer. During the intervention phase, applicants were initially shortlisted on the basis of their de-identified CVs. Once initial shortlisting decisions had been made, applicant information was reinstated, and shortlisting decisions were re-evaluated prior to moving on to the interview stage (see Figure y.) We therefore compare the gender and university composition of applicants shortlisted during the baseline phase (de-identification), with the demographic composition of applicants on the initial and revised shortlists during the intervention phase (Outcome Measures 1 and 2 in Figure y.).

Analyses

Two research questions were analysed using a separate moderated linear regression model. The model examined the direct effect of the predictor variable (gender or university attended) on the recruitment outcome variable. The model also examined whether any direct relationships between predictor variables and the recruitment outcome were moderated by CV de-identification. The model was run twice for each predictor, initially without covariates, and a second time including covariates.

The two research questions were:

1. Does de-identification of gender reduce the hiring gap between men and women during recruitment?
2. Does de-identification of higher education institution improve recruitment equality for applicants from institutions of varying prestige?

Results

Demographics

Table 8 displays participant demographic data. Overall, there was a higher number of male than female applications, but with similar proportions of males and females across the two conditions. There was a larger proportion of participants from Go8 universities in the baseline condition than in the de-identified condition.

Table 8
Participant demographic data for the overall DTF sample

| | | Baseline | De-identified |
|------------|--------------------|------------|---------------|
| | Total participants | 42 | 120 |
| | Male | 18 | 49 |
| | Female | 10 | 34 |
| | Unspecified Gender | 28 | 37 |
| University | Go8 | 31 (73.8%) | 69 (57.5%) |
| | Non-Go8 | 9 (21.4%) | 49 (40.8%) |
| | Missing | 2 | 2 |

Correlations

Pearson correlations between variables revealed a significant negative correlation between attending a Go8 university and being in the de-identified experimental condition. No other correlations between predictor and outcome variables met significance (see Table 9.)

Table 9

Pearson correlations between participant gender, Go8 university attended or not, experimental condition, and application stage reached.

| Variables | Mean | SD | 1 | 2 | 3 | 4 | 5 |
|--|------|-----|------|-------|------|---|---|
| 1. Participant Gender | .60 | .49 | | | | | |
| 2. Go8 University | .63 | .48 | .02 | | | | |
| 3. Baseline or De-Identified Condition | .74 | .44 | -.05 | -.17* | | | |
| 4. Application Stage Reached | .40 | .64 | .10 | .12 | -.03 | | |

Notes. Gender: female = 0, male = 1; 0 = Non-Go8, 1 = Go8; 0 = Baseline, 1 = De-identified; Application stages include 0 = Not shortlisted, 1 = Interview, and 3 = Job offer. N = 110-162. * $p < .05$, ** $p < .01$.

Research Question 1

Does de-identification of gender reduce the hiring gap between men and women during recruitment?

Research Question 1 was analysed using a moderated linear regression model. The model examined the direct effect of gender on recruitment outcome, and whether this relationship was moderated by the intervention (i.e., baseline or de-identification condition). The model was also run a second time to include the additional predictor variable (university attended) as a covariate.

The overall linear regression model was not significant, $F(3, 107) = 2.07, p = .108$. However, results indicate a main effect of gender on recruitment outcome, $t(107) = 2.29, p = .024$, with males more likely to proceed further through the recruitment process than females in the baseline condition. There was also an interaction between gender and de-identification to predict recruitment outcome, $t(107) = -2.06, p = .042$, whereby the hiring gap between men and women was no longer present following de-identification. When university was added as a covariate, these two effects further increased in magnitude. The relative percentages of male and female applications at different stages of recruitment, across conditions, is shown in Figure 10.

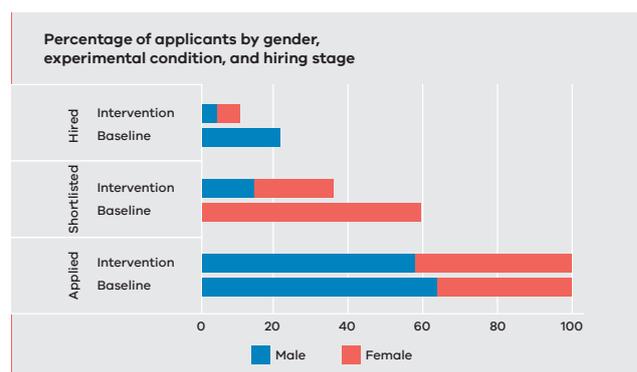


Figure 10. Percentage of total male and total female applicants who applied, were shortlisted, or hired during baseline and de-identification phases, relative to the number of applicants in total within the gender category.

In addition, DTF undertook a specific form of de-identification whereby applicants were de-identified in the first instance, then were re-identified before a final decision was made regarding shortlisting for interview. We undertook additional exploratory analyses to better understand the impact of this design. Figure 11 displays the results of this process, showing that de-identification improves women's likelihood of being shortlisted, and that this progress continues even after the gender of the applicant is revealed before the final shortlist is confirmed.

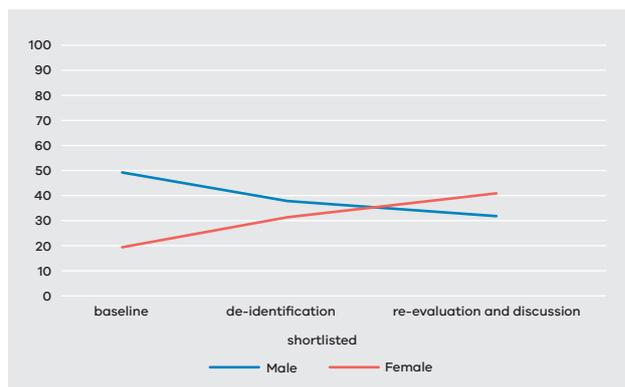


Figure 11. Percentage of female applicants being shortlisted, both under conditions of de-identification and after re-identification.

Research Question 2

Does de-identification of higher education institution improve recruitment equality for applicants from institutions of varying prestige?

Research Question 2 was analysed using a moderated linear regression model. The model examined the direct effect of candidate university attended (Go8 versus other), and whether this relationship was moderated by the intervention (i.e., baseline versus de-identified condition). The model was also run a second time to include gender as a covariate.

The overall linear regression model was significant, $F(3, 154) = 3.45, p = .018$. Results indicated a main effect of attending a Go8 University on recruitment outcome that approached significance, $t(154) = -1.90, p = .059$, with candidates from non-Go8 universities progressing further through the recruitment process than their Go8 counterparts in the baseline condition. De-identification significantly moderated the relationship between university and recruitment outcome, $t(154) = 2.84, p = .005$, with de-identification now resulting in Go8 applicants proceeding further through the recruitment process than their non-Go8 counterparts. There was little difference in the model when the gender covariate was included. This finding is shown in Figure 12.

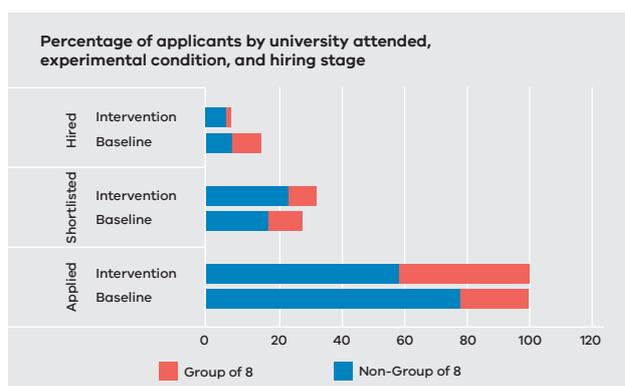


Figure 12. Percentage of total Group of 8 and non-Group of 8 applicants who applied, were shortlisted, or hired during baseline and de-identification phases, relative to the number of applicants in total within the university category.

Discussion

We tested whether CV de-identification affected how far diverse candidates progressed through the recruitment process. The variables de-identified included candidate gender and university attended. We found that before de-identification, male applicants enjoyed an advantage in the hiring process which was eliminated under conditions of de-identification. However, we found an unexpected result with higher education. Before de-identification, applicants from non-Go8 universities outperformed Go8 graduates on hiring, however this was reversed after de-identification, with applicants from Go8 universities progressing further through hiring.

The finding that de-identification assists female applicants in progressing through recruitment was in line with our hypotheses. This finding also corresponds with prior research that suggests anonymising CVs can increase interviews and job offers for female candidates (Åslund & Nordström Skans, 2007; Behaghel et al., 2012). The specific role targeted for this recruitment pilot was also a male-dominated one, reflected in the gender composition of applicants – two thirds of applicants were men. Nonetheless, de-identification of gender suggests that women are equally qualified for the role and that assumptions of gender may be interfering with the objectivity of the hiring process. However, this pilot contradicts a similar study conducted in a European research institution, which found that female PhD graduates in a masculine-gendered role were disadvantaged by the de-identification process, possibly as a result of positive discrimination being eliminated (Krause, Rinne, & Zimmermann, 2012).

We also found a significant effect of de-identifying applicant's higher education institution; however, this finding was in the opposite direction to our hypothesis. Graduates from Go8 institutions progressed further through hiring after rather than before the de-identification process. The reasons for this reversal of effects is unknown, but three possible explanations are suggested for these findings. First, there may be a genuine qualitative difference between the quality of graduates from different universities.

Future field studies may consider including an objective indicator of educational attainment, such as GPA, to test for whether this may affect hiring biases. Second, the non-Go8 sample may have been biased by applicants from high-ranking international institutions, while our coding only accounted for perceived prestige of national institutions. However, this is unlikely given that far more applicants were from non-Go8 universities in the de-identified dataset, yet it was more likely for Go8 applicants to progress through hiring in that condition. Third, we categorised higher education according to whether the applicant had ever attended a Go8 university, rather than by their most recent educational qualification. This may have also impacted the results.

Overall, CV de-identification appears to be of particular importance for gender. A preference for male candidates is eliminated under conditions of de-identification. However, we find a contradictory result for higher education, with de-identification improving the hiring progress of applicants from prestigious institutions i.e., Go8 universities.

Case Study on CV De-Identification

Department of Justice and Regulation

Participants

The sample for this dataset comprised 151 applicants (88 female, 62 male, 1 unspecified) who applied for one of 9 roles within the Department of Justice and Regulation (DJR) between February and August 2017. The baseline condition (from 23 February to 30 March, 2017) comprised 53 applicants, and there were 98 applicants in the de-identified condition (April to August, 2017). Demographic variables are summarised in Table X. Complete data for gender was obtained for $n = 150$ participants, with $n = 133$ participants providing suburb data and $n = 123$ participants providing university attended data. As such, some of the analyses contain samples smaller than the overall number of participants in this project.

Study design

The study included measures of recruitment outcomes prior to and following the introduction of CV de-identification. In the de-identified condition, all hiring managers made shortlisting decisions based on the de-identified version of the applicants' CV.

Procedures and measures

The intervention used an online recruitment tool to de-identify certain demographic information relating to candidates during the hiring process. Baseline data with standard identification was collected in February and March 2017. From April to August 2017, the following information was de-identified in all applications: applicant name, applicant age, applicant gender, applicant home address, applicant university/institution, and applicant secondary school attended.

Predictor variables

We conducted evaluations on the following three variables of interest: applicant gender, applicant home address, and applicant university attended. The latter two variables were taken to reflect social capital.

Applicants were coded as male or female based on their self-selected title (e.g., Mr, Ms). Where applicants used a non-gendered title, such as Dr, gender was coded as missing data and excluded from gender analyses. Home address (suburb only) was coded from 1 to 10 using the Australia Bureau of Statistics' (2018) Index of Relative Socio-Economic Advantage and Disadvantage (IRSAD). A low score (minimum = 1) indicates relatively greater disadvantage and a lack of advantage in general, while a high score (maximum = 10) indicates a relative lack of

disadvantage and greater advantage in general. Higher Education institution attended was coded as a binary variable based on whether the institution attended by the candidate was a member of the Group of Eight⁷ (Go8; considered a marker of prestige), or otherwise. Across government roles, alumni of Go8 institutions – especially the University of Melbourne and Monash University – appear to enjoy preferential hiring outcomes. We thus wanted to investigate whether de-identifying this information influenced recruitment progress for applicants from other tertiary institutions.

Moderator variable

CV de-identification was integrated into the existing DJR hiring process. In DJR’s standard hiring practice, applicants who submit CVs for a role are reviewed by hiring managers who shortlist a number of candidates for interview. After the interview, referee checks are conducted and successful applicants are offered roles. De-identification was in place for CVs between the application and shortlisting for interview stage.

Outcome variable

The outcome of interest was how far candidates proceeded through the recruitment process. This was measured on a four-point scale that included Application Stage, Reference Check, Interview, and Job Offer.

Analyses

Three research questions were analysed using separate moderated linear regression models. Each model examined the direct effect of the predictor variable (gender, suburb, or university attended) on the recruitment outcome variable. The model also examined whether any direct relationships between predictor variables and the recruitment outcome were moderated by CV de-identification. The model was run twice for each predictor, initially without covariates, and a second time including covariates.

The three research questions were:

1. Does de-identification of gender reduce the hiring gap between men and women during recruitment?
2. Does removing home addresses from CVs equalise recruitment outcomes for applicants from socio-economically diverse suburbs?
3. Does de-identification of higher education institution improve recruitment equality for applicants from institutions of varying prestige?

Table 10
Participant demographic data for the overall DJR sample

| | | Baseline | De-identified |
|------------|---------------------|-------------|---------------|
| | Total participants | 53 | 98 |
| | Male | 16 (30.8%) | 46 (46.9%) |
| | Female | 36 (67.9%) | 52 (53.1%) |
| | Unspecified Gender | 1 | 0 |
| Suburb | Average IRSAD score | 7.60 (3.04) | 8.51 (2.07%) |
| | Not Identifiable | 8 (15.1%) | 10 (10.2%) |
| University | Go8 | 19 (35.8%) | 43 (43.9%) |
| | Non-Go8 | 22 (41.5%) | 39 (39.8%) |
| | Missing | 12 (22.6%) | 16 (16.3%) |

7: Go8 universities include The University of Adelaide, The Australian National University, The University of Melbourne, Monash University, The University of New South Wales, The University of Queensland, The University of Sydney, and The University of Western Australia

Results

Demographics

Table 10 displays participant demographic data. Overall, there was a slightly greater number of female applicants, with a slighter higher proportion of males in the de-identified compared to the baseline condition. Suburb relative advantage to disadvantage was marginally higher for participants in the de-identified group, while there was a similar proportion of participants from Go8 and non-Go8 universities across the two conditions.

Table 11

Pearson correlations between participant gender, suburb IRSAD score, Go8 University attended or not, experimental condition, and application stage reached.

| Variables | Mean | SD | 1 | 2 | 3 | 4 | 5 |
|--|------|------|-------|------|-----|------|---|
| 1. Participant gender | .41 | .49 | | | | | |
| 2. Suburb | 8.20 | 2.47 | .04 | | | | |
| 3. Go8 University | .50 | .50 | -.20* | .27* | | | |
| 4. Baseline or De-Identified Condition | .65 | .48 | .16 | .18* | .06 | | |
| 5. Application Stage Reached | .44 | .85 | .002 | -.03 | .02 | -.06 | |

Notes. Gender: female = 0, male = 1; 0 = Non-Go8, 1 = Go8; 0 = Baseline, 1 = De-identified; Application stages include 0 = Application, 1 = Reference Check, 2 = Interview, and 3 = Job Offer. N = 109-151. * $p < .05$, ** $p < .01$.

Research Question 1

Does de-identification of gender reduce the hiring gap between men and women during recruitment?

Research Question 1 was analysed using a moderated linear regression model. The model examined the direct effect of gender on recruitment outcome, and whether this relationship was moderated by the intervention (i.e., baseline or de-identification condition). The model was also run a second time to include the additional two predictor variables (applicant home address and university attended) as covariates.

The overall linear regression model was not significant, $F(3, 146) = .46, p = .713$. Results indicate no main effect of gender on recruitment outcomes, $t(146) = -.70, p = .49$, and no interaction between gender and de-identification to predict recruitment outcomes, $t(146) = .93, p = .352$. There was little difference in the model when covariates were included.

Correlations

Pearson correlations between variables revealed significant negative correlations between attending a Go8 university and male gender. Significant positive correlations were present between attending a Go8 university and living in a more socio-economically advantaged suburb, and between being in the de-identified experimental condition and living in a more socio-economically advantaged suburb (see Table 11).

Research Question 2

Does removing home addresses from CVs equalise recruitment outcomes for applicants from socio-economically diverse suburbs?

Research Question 2 was analysed using a moderated linear regression model. The model examined the direct effect of candidate suburb on recruitment outcomes, and whether this relationship was moderated by the intervention (i.e., baseline or de-identification condition). The model was also run a second time to include covariates. These included participant gender, and whether the participant attended a Go8 university.

The overall linear regression model was not significant, $F(3, 129) = .41, p = .746$. Results indicated no main effect of candidate suburb on recruitment outcomes, $t(129) = .41, p = .684$, and no interaction between de-identification and suburb to predict recruitment outcomes, $t(129) = -.74, p = .463$. There was little difference in the model when covariates were included.

Research Question 3

Does de-identification of higher education institution improve recruitment equality for applicants from institutions of varying prestige?

Research Question 3 was analysed using a moderated linear regression model. The model examined the direct effect of candidate University attended (Go8 versus other), and whether this relationship was moderated by the intervention (i.e., baseline or de-identification condition). The model was also run a second time to include covariates. These included participant gender, and participant home suburb socio-economic status.

The overall linear regression model was not significant, $F(3, 119) = .97, p = .408$. Results indicated a no main effect of attending a Go8 university on recruitment outcomes, $t(119) = .88, p = .379$, and no interaction between de-identification and Go8 university and recruitment outcomes, $t(119) = -.84, p = .402$. There was little difference in the model when covariates were included.

Discussion

We tested whether CV de-identification affected how far diverse candidates progressed through the recruitment process. Overall, we found no effect for any of our variables of interest: gender, applicant home suburb, or higher education institution. There was no preference for one gender of applicant over another, nor was there a preference for applicants who live in suburbs with higher socioeconomic status. We also did not find a preference for applicants from more prestigious higher education institutions and de-identifying this information did not have a significant impact on progression through the recruitment process.

The lack of significant findings for this dataset may be the result of a relatively small sample per role, with only 151 applicants spread across 9 roles. Given the reduction in applicants at each stage of the recruitment process, we may not have had large enough numbers to identify trends and changes in the diversity composition of applicants. A larger sample from a smaller set of roles may assist in pinpointing where CV de-identification can affect recruitment outcomes. Furthermore, similar to the Department of Premier and Cabinet and unlike the Department of Treasury and Finance, the Department of Justice and Regulation undertook this pilot across a diverse range of roles ranging in required skills and qualifications. A focus on roles that currently lack diversity of staff may be a more beneficial design for conducting pilot testing of CV de-identification in this organisation.

Finally, similar to the Department of Premier and Cabinet, the Department of Justice and Regulation have attracted applicants who predominantly live in suburbs with high socioeconomic status. We recommend considering different means of attracting applicants from a greater diversity of socioeconomic backgrounds to ensure greater representation of the Victorian population.

Overall, CV de-identification did not significantly impact the diversity of applicants for the Department of Justice and Regulation. However, results also suggest that the Department of Justice and Regulation is currently not preferencing applicants of specific genders, socioeconomic status, or status of educational institution.

General Discussion of CV De-Identification

Five organisations participated in the CV de-identification intervention. This involved removing a range of demographic characteristics from application forms, in order to investigate whether this information influenced candidate recruitment outcomes. Examples of information that was de-identified include applicant name, gender, country of birth, higher education institution, and home address. The results of this pilot program indicated that a number of recruitment decisions were improved by this intervention. CV de-identification resulted in significant improvements to recruitment of applicants who were born in an overseas country (VicRoads), who lived in lower socioeconomic suburbs (Department of Premier and Cabinet), and whose gender was under-represented in the targeted role (Department of Treasury and Finance). These results suggest that CV de-identification can increase access and equity of outcomes for applicants from specific social groups.

Recruitment outcomes prior to de-identification indicated inequality in recruitment based on country of birth. Overseas-born applicants were less likely to be shortlisted for roles, despite having more years of experience compared to Australian-born applicants. Removing country of birth from applications reversed the association completely, such that Australian-born applicants were preferred by hiring managers who received identified CVs, while overseas-born applicants were preferred by hiring managers who received de-identified CVs. This was likely driven by a difference in the average level of experience. Overseas-born applicants typically had more

experience in the role compared to Australian-born candidates. When the country of birth was de-identified, years of experience may have been weighed more heavily in the shortlisting decision.

While CV de-identification led to a reversal of preferences regarding applicant's country of birth, it levelled the playing field for other demographics. This was the case at the Department of Treasury and Finance, who obscured the gender of applicants for traditionally male-dominated roles. When gender was identified, male applicants were preferred for the role. After de-identification, neither gender was preferred. Similarly, removing home address from applications improved the progression of applicants from suburbs with lower socioeconomic status at the Department of Premier and Cabinet.

De-identification of gender results in equalising recruitment of men and women in organisations and roles where an existing gender imbalance in hiring was noted, such as at the Department of Treasury and Finance. However, de-identification of gender did not result in differences in the recruitment ratio of men and women when organisations did not have a pre-existing gender inequality in their recruitment outcomes, as was the case at Hall & Wilcox, Department of Premier and Cabinet, and Department of Justice and Regulation. Given the relatively equal progress of applicants in these roles, CV de-identification was unlikely to change this dynamic as there was no gender bias to correct. However, in the case of VicRoads, there were unequal hiring outcomes due to a severe imbalance in the

number of female applicants; in this case, there were not enough female candidates in the applicant pool to detect a significant change in shortlisting outcomes. In cases like this, CV de-identification must be paired with additional measures to improve the pipeline of applicants from the underrepresented minority group.

De-identifying home addresses improved recruitment progress for applicants from lower socioeconomic suburbs for applicants at the Department of Premier and Cabinet, but not at the Department of Justice and Regulation. Possible reasons for this contrast in findings may be related to the specific roles being recruited for. For example, the majority of roles at the Department of Justice and Regulation were roles requiring a law degree or legal background (e.g., senior legal policy officer), while roles at the Department of Premier and Cabinet were typically more generalist policy or project roles. In both organisations, applicants from lower socioeconomic suburbs were strongly outnumbered by those from higher socioeconomic suburbs. Given that the government aims to represent all its constituents, we recommend that both departments consider how they may better attract applicants from a diverse range of socioeconomic backgrounds to apply for roles. This may involve advertising roles in targeted locations, attending local events in diverse suburbs, providing representation at employment fairs across different educational institutions, etc.

Furthermore, some of the trial results showed that while bias was occurring in the recruitment process, de-identification did not necessarily improve equity for applicants. This included evidence of ethnicity bias towards Western applicants for Hall & Wilcox, and a bias towards more prestigious higher education institutions at the Department of Premier and Cabinet. In these cases, a lack of change resulting from CV de-identification may illustrate a skills or experience gap between applicants from different social groups. However, it may also arise from some applicants lacking a robust understanding of unspoken social norms and implicit knowledge about recruitment. These applicants may be equally qualified for roles, but lack the presentation needed to convey their skills and abilities. Future research may examine which of these hypotheses best explains these recruitment trends. In the latter case, organisations wishing to recruit a more diverse workforce may want to consider how to best support minority group candidates with their application to ensure that their skills and expertise are appropriately conveyed, to improve success rates in recruitment.

Finally, it is also important to note that while these organisations all implemented some form of CV de-identification, there were a range of differences between interventions that complicate direct comparison of results across organisations. Some organisations chose to focus on specific roles, such as seasonal law clerks or project engineers, while other departments conducted de-identification across a broad range of roles within a specified time frame, resulting in a highly variable set of roles being tested. Organisations also ranged on the size of the dataset obtained, from a relatively small dataset of 151 applicants to a much larger sample of over 1,500 applicants. We recommend caution when interpreting the results of smaller datasets, as the findings may be less stable than those of larger datasets.

Overall, this set of CV de-identification trials show promising results for improving equity in recruitment of different social groups. We caution that this form of intervention may be best suited to roles that have relatively strong representation of minority social groups applying for roles, and that also show a current inequality in the selection of applicants. For roles with limited diversity in the applicant pool, interventions should be made earlier in the process, such as rewording of the job advertisement, to ensure that the applicant pool contains a diverse range of applicants from which to select, before de-identification is undertaken. Organisations that gained limited benefit from CV de-identification were typically those that already had relatively good equity of access for select social groups. In those cases, alternative diversity strategies may be employed, such as unconscious bias training or quotas in leadership roles for minority applicants.

Intervention

4

Unconscious Bias Training

Unconscious bias, defined as subtle subconscious beliefs and opinions that influence people's behaviour, remains a persistent issue in the recruitment and promotion process.

Consequences of unconscious bias include, workplace conflicts, inefficiency, inappropriate management decisions, and employee dissatisfaction.

Unconscious bias training has emerged as one of several attempts to curtail the negative impacts of bias arising from demographic and social differences between people. This training helps participants better identify situations in which their decision making may be influenced by unconscious knowledge, and proactively respond.

Training context and design

Training programs are most effective when tailored to the unique needs and characteristics of the target audience. A bias training program conducted as one of many programs that make up a larger diversity initiative across the organisation will be more effective than a standalone seminar (Kalev et al., 2006). There is no evidence that unconscious bias training by itself leads to greater workforce diversity (Kalev et al., 2006). However, there is support for the claim that unconscious bias training leads to reductions in implicit biases and behavioural change (Forscher, Mitamura, Dix, Cox, & Devine, 2017). Forscher et al., (2017) showed that unconscious bias training produced effective changes in awareness, implicit bias and a measures of voluntary diversity behaviour taken two years after the training.

Multiple reviews have identified features of diversity training programs, including unconscious bias training, that lead to more positive effects for raising awareness, reducing bias, changing attitudes and influencing behaviours (Bezrukova, Spell, Perry, & Jehn, 2016; Mollica & Friedman, 2007; Paluck & Green, 2009). Mandatory training appears to be more effective, whereas voluntary training is perceived more favourably by training participants.

This is likely because people who choose to participate are more likely to go in with favourable dispositions to diversity and unconscious bias training principles. Justifications that stress norms of improved decision making and interpersonal relationships are more effective than an emphasis on 'everybody is biased'. Effective content has been found to include; concepts and examples of bias; exercises and role plays; an implicit attitude (unconscious knowledge) measure and feedback; and, development of strategies and action plans for implementation on the job.

Longer, more comprehensive programs tend to also be more effective than shorter programs. Longer programs provide more opportunities for contact, practice, and reinforcement of new learning. Diversifying the training approach by incorporating a combination of methods (e.g., lectures, exercises, activities, video presentations) increases participants' enjoyment of the process, increases likelihood of a favourable response, and maximises opportunities to learn (Kalinowski, Steele-Johnson, Peyton, Leas, Steinke, & Bowling, 2013).

Finally, by increasing awareness of unconscious bias and the associated risks, participants are better able to identify these situations and proactively respond. UBT is most effective when the program has discrete, comprehensible objectives. Broad, sweeping, generalist trainings often do little to engage participants, and result in little actual improvement. A considered, well planned program aimed at improving specific employee behaviours and interactions, on the other hand, has a much greater chance of success.

Risks and barriers to successful outcomes

Increased awareness does not automatically lead to behavioural change, as people do not always know how to improve their biases. Participants may also be resistant to training if it does not align with their values or personal beliefs. Some time may be required for participants to develop and adapt their behaviour in line with the information learned about diversity. Measuring success indicators across multiple time points may thus show changes over time more successfully.

There is a risk that programs can leave those not in the targeted minorities feeling guilty and 'cast as oppressors'. Such friction can cause division amongst the workplace. This can largely be mitigated through the structure, design and approach of the training, so that it is cooperative rather than combative in nature. Unconscious bias training also runs a risk of giving trainees that they are 'more rational' and, because they are now aware of their biases, that the existing system is fair.

Measuring outcomes

Effective unconscious bias training programs can result in greater awareness of bias, reduced implicit bias, more positive attitudes towards diversity and a greater willingness to engage in diversity related behaviours. Training effectiveness and success can be measured via a combination of self-report and observational measures. For example:

- Participants may complete a survey before and after undergoing training, with information about subconscious biases and prejudices, as well as their general attitude and behaviours, compared across time points.
- Participants can be directly interviewed regarding their perceptions and opinions of the impact of training.
- Participants' behavioural and attitudinal changes can be observed by others. Successful training will likely result in observable social and conventional changes in the workplace, and a larger trend to a more inclusive, bias-free environment.

A key measure of training success will be the endurance of training impact over time, with participants displaying behaviour and attitudes that correspond with the training principles over a sustained period of time.

Case Study

Unconscious Bias Training across Victorian Government Departments

Participants

Participants were 221 individuals sampled from seven Victorian Government departments or affiliates who participated in this intervention: Department of Premier and Cabinet (DPC); Department of Treasury and Finance (DTF); Emergency Management Victoria (EMV); VicRoads; Ambulance Victoria (AV); Transport Accident Commission (TAC); and the Victorian

Equal Opportunity and Human Rights Commission (VEOHRC). Each organisation nominated up to 100 employees to participate in unconscious bias training. Participants were predominantly, but not exclusively, drawn from areas associated with hiring and promotion.

Table 12 lists the number of participants from each body at each stage of the intervention. Only participants who completed each stage moved on to the next step of the program.

Table 12

Number of participants from each organisation at each stage of the unconscious bias program

| | DPC | DTF | EMV | Vic Roads | TAC | VEOH RC | AV | Total |
|--|-----|-----|-----|-----------|-----|---------|----|-------|
| Nominated | 100 | 101 | 111 | 108 | 30 | 55 | 63 | 568 |
| Completed Survey Time 1 | 61 | 82 | 85 | 46 | 23 | 35 | 39 | 371 |
| Completed Survey Time 2 (intervention) | 10 | 21 | 27 | 5 | 6 | 6 | 12 | 87 |
| Completed Survey Time 2 (control) | 25 | 27 | 37 | 13 | 4 | 13 | 15 | 134 |
| Final sample | 35 | 48 | 64 | 18 | 10 | 18 | 27 | 221 |

Study design

The study involved taking measures of attitudes to diversity, organisational fairness, psychological safety, self-efficacy, behavioural intentions, and self-reported behaviours prior to and following the unconscious bias training modules.

Procedures and measures

Participants were initially administered the first of two surveys relating to workplace diversity, which was administered online and took 10 to 15 minutes to complete. It measured the five variables that were formed via factor analysis. While the survey was initially conceived as comprising six factors, reliability analyses suggested that self-efficacy and behavioural intentions were best collapsed into a single factor. The final factor structure is displayed in Appendix B, with correlations and reliability displayed in Table 15. The variables included:

Attitudes to diversity was measured by five items on a five-point scale, ranging from 1 *Strongly disagree* to 5 *Strongly agree*; e.g., "I consider lack of diversity at work to be a serious social problem", "I am not personally concerned about lack of diversity"

Organisational fairness was also measured by five items on a five point scale, ranging from 1 *Strongly disagree* to 5 *Strongly agree*; e.g., "In your organisation or industry, people are hired because of their skills and abilities", "In your organisation, opportunities are distributed evenly"

Psychological Safety was measured by four items on a seven-point scale, ranging from 1 *Very inaccurate*, to 7 *Very accurate*; e.g., "I am able to bring up problems and tough issues in my organisation", "It is safe to take a risk in this organisation"

Self-efficacy and behavioural intentions was measured by 15 items on a 5 point scale, ranging from 1 *Cannot do at all or Will not engage at all*, to 5 *Certainly can do or Will definitely engage*; e.g., "I am certain I can advocate for diversity in this organisation on a regular basis", "I intend to speak up and challenge discriminatory humour in the next six months"

Reported behaviour was measured by participants reporting how many behaviours, out of a potential ten, they had engaged in over the past 6 months, or since completing the first survey. Examples of behaviours included initiating a conversation about diversity and inclusion with a colleague at work or a supervisor, or intervening when witnessing discrimination or bullying.

After administration of the survey, participants were invited to attend a one-hour workshop conducted by Professor Robert Wood. The workshop outlined the theoretical basis for the training, typical results, and answered questions frequently asked by training attendees. The goal of the workshop was to maximise commitment to the training program. Participants then engaged in a cross-over design. The full sample was split into two, with one group (intervention group) receiving training earlier than the other group (control group). Allocation to training groups was random, but a limited number of scheduling conflicts were taken into account in group allocation. A second survey was administered after the intervention group had undergone training, but before the control group had received their training. This survey was identical, with the exception of wording being adjusted to account for the passage of time (i.e. items such as 'have you done this in the last six months' were replaced with 'have you done this since the last survey'). Figure 13 illustrates this cross over design.

The online training program was provided by Cognicity (<http://www.cognicity.com>) and comprised three online learning modules, including an implicit association measure of unconscious knowledge. The implicit association task is based on the Go/No go Association Task (GNAT, Nosek & Banaji, 2001) and provides, for each individual, a measure of their unconscious associations between gender and leadership, and gender and domesticity. Participants completed three 30-minute modules. The content provided an evidence-based understanding of unconscious knowledge and bias, exercises, and strategies for recognising and reducing unconscious bias in the workplace. The content of the three modules covered three themes:

- How people think;
- Unconscious bias, including the implicit association task and feedback report;
- Strategies for mitigating bias.

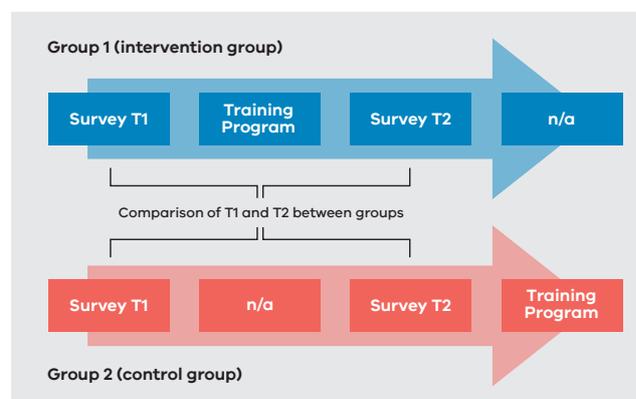


Figure 13. Cross-over design showing relative time points for training and data collection.

Analyses

In order to test the effect of the unconscious bias training intervention, a 2 (time: Time 1, Time 2) by 2 (group: control, intervention) mixed model ANOVA was conducted, with separate models testing each outcome variable measured by the survey.

The two research questions were:

1. Does participating in unconscious bias training increase positive attitudes toward diversity?
2. Does unconscious bias training improve the likelihood of participants engaging in diversity-supporting behaviours, via their increased self-efficacy and intentions?

Table 13

Range and Mean Response for Age, Years in Job and Number of People Supervised

| | Minimum | Maximum | Mean |
|-----------------------|---------|---------|-------|
| Age | 24 | 66 | 43 |
| Years in job | 1 | 18 | 3 |
| No. people supervised | 0 | 750 | 14.37 |

Table 14

Gender, English Language, Country of Birth, and Education Level of Participants

| Variable | Number of Participants (% in brackets) |
|----------------------------------|--|
| Gender | |
| Male participants | 112 (48%) |
| Female participants | 0106 (50.7%) |
| Not specified | 3 (1.4%) |
| English as First Language | |
| Yes | 209 (94.6%) |
| No | 9 (4.1%) |
| Country of Birth | |
| Australia | 183 (82.8%) |
| Elsewhere | 35 (15.8%) |
| Education | |
| Did not complete high school | 7 (3.2%) |
| Diploma or certificate | 26 (11.8%) |
| Undergraduate degree | 26 (11.8%) |
| Postgraduate degree | 104 (47.1%) |

Results

Demographics

Tables 13 and 14 display participant demographic data. Overall, there was an even number of male and females, with most participants born in Australia with English as their first language and holding a post-graduate qualification.

Correlations

Pearson correlations were conducted between the following variables at Time 1 and Time 2: Self-reported behaviours, attitudes to diversity, organisational fairness, psychological safety, and efficacy-intentions. There were significant positive correlations between the time 1 and time 2 variables (e.g., Behaviour at Time 1 and Time 2).

Other significant correlations included positive correlations between psychological safety, organisational fairness, and efficacy-intentions. Notably, there was a negative correlation between perceptions of organisational fairness and self-reported diversity behaviours (see Table 15.)

Table 15
Correlations between Training, Behaviour, Attitudes, Efficacy-intentions, Organisational Fairness and Psychological Safety at Time 1 and Time 2

| Variables | Mean | SD | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---------------------|------|------|-------|--------|--------|-------|--------|-------|-------|-------|-------|-------|-----|
| 1. Training | | | | | | | | | | | | | |
| 2. T1 Behaviour | 3.05 | 2.04 | -.06 | .65 | | | | | | | | | |
| 3. T2 Behaviour | 3.71 | 2.14 | .03 | .47** | .65 | | | | | | | | |
| 4. T1 Attitudes | 3.97 | 0.75 | -.05 | .27** | .30** | .75 | | | | | | | |
| 5. T2 Attitudes | 4.11 | 0.70 | .05 | .28** | .40** | .72** | .78 | | | | | | |
| 6. T1 OrgFairness | 4.47 | 0.35 | -.03 | -.18** | -.17** | -.06 | -.04 | .74 | | | | | |
| 7. T2 OrgFairness | 4.57 | 0.34 | .04 | -.16* | -.09 | -.03 | -.09 | .66** | .78 | | | | |
| 8. T1 PsychSafety | 3.28 | 0.75 | .03 | -.06 | -.02 | .02 | .01 | .49** | .44** | .72 | | | |
| 9. T2 PsychSafety | 3.31 | 0.76 | .09 | -.06 | .01 | .00 | .01 | .44** | .53** | .66** | .66 | | |
| 10. T1 Effic-Intent | 4.76 | 1.23 | .01 | .11 | .10 | .24** | .22** | .10 | .05 | .21** | .15* | .80 | |
| 11. T2 Effic-Intent | 4.93 | 1.13 | .02** | .12 | .23** | .26** | .39*** | .05 | .15* | .27** | .25** | .58** | .83 |

Notes. Cronbach's alphas are listed on the diagonal. * $p < .05$, ** $p < .01$.

Research Question 1

Does participating in unconscious bias training increase attitudes and behaviours supportive of diversity?

First, participants in the control and experimental groups recorded no difference in any of the outcome variables at baseline indicating that the random allocation process produced equivalent groups. In terms of changes in participant responses from the first survey to the second survey, we saw a number of positive impacts on diversity, regardless of whether participants had received training or not. We also saw change in the efficacy-intention variable. For participants in the control and experimental groups, between Time 1 and Time 2, we noted the following findings:

Attitudes: Participants reported more positive attitudes toward at Time 2 compared to Time 1 $F(1,219) = 16.62, p < .001$. This increase was evident for participants in the experimental and control groups, $F(1,219) = .001, p = .976$, although there was a trend for participants in the experimental group to show marginally more positive attitudes at Time 2 compared to individuals in the control group, $F(1,219) = 3.68, p = .057$.

Past Behaviour: Participants also reported engaging in a higher number of behaviours supportive of diversity between Time 1 and Time 2, $F(1,219) = 22.62, p < .001$. This increase again was evident for participants in both groups, $F(1,219) = .03, p = .855$.

Organisational Fairness: There was no change in perceptions of organisational fairness between Time 1 and Time 2, and no difference between those in the experimental and control groups.

Psychological Safety: Participants reported greater feelings of psychological safety at Time 2 compared to Time 1, $F(1,219) = 7.58, p = .006$. This increase was also evident for participants in the control and experimental groups, $F(1,219) = 1.12, p = .290$.

Efficacy-Intentions: Participants reported feeling greater self-efficacy and intending to act in diversity-friendly ways at Time 2 compared to Time 1, $F(1,219) = 26.73, p < .001$. This was qualified by a significant interaction with group, $F(1,219) = 10.13, p = .002$. Simple effects analyses show that participants in the intervention group reported significantly higher efficacy-intentions at Time 2 (after training: $M = 4.66, SD = 0.04$), than those in the control group ($M = 4.52, SD = 0.03$) who had not received training at Time 2, $F(1,219) = 9.49, p = .002$, see Figure 14.

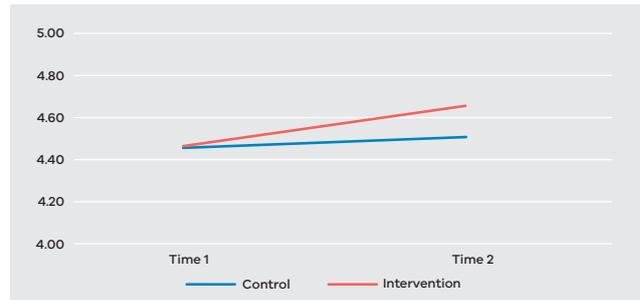


Figure 14. Impact of unconscious bias training on self-reported participant efficacy and intentions. Only the intervention group received training prior to survey Time 2.

Research Question 2

Does unconscious bias training improve the likelihood of participants engaging in diversity-supporting behaviours, via their increased self-efficacy and intentions?

A large body of research evidence has demonstrated strong causal relationships between self-efficacy and intentions, and reported behaviour (Bandura, 1997). Thus, we undertook a mediation analysis to explore the link between unconscious bias training, increases in efficacy-intentions, and increases in reported diversity behaviours. The mediation analysis was conducted using Hayes PROCESS Model 4 in SPSS, with a bootstrap of 10,000 and 95% confidence intervals. The full model and results are shown in Figure 15.

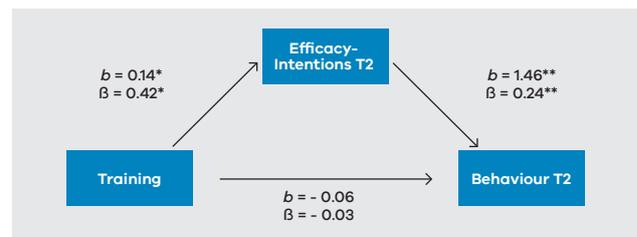


Figure 15. Standardised and unstandardized regression coefficients for the relationship between unconscious bias training (control group = 1, intervention group = 2) and behaviour measured at time 2, as mediated by efficacy and intentions at time 2. * $p < .01$, ** $p < .001$

Results show no direct effect of training on behaviours at Time 2. However, there was an indirect effect of training on behaviour, via the effect of training on participant efficacy and intentions, which subsequently predicted T2 behaviour, $b = .21, 95\%$ Confidence Interval (.06, .39). The positive impact of unconscious bias training on the frequency of diversity related behaviours on the job only occurred if the training increased participants efficacy-intentions for implementing those behaviours.

Discussion

We tested the effectiveness of unconscious bias training on a sample of participants from seven Victorian government organisations and affiliates.

Overall, we saw increases in responses to diversity-supportive attitudes and behaviours across the board from Time 1 to Time 2. However, the training program augmented improvements on two key variables; participants in the experimental group showed higher levels efficacy-intentions. That is, as a result of the training, participants felt more confident about their ability to engage in diversity-supporting behaviours and also intended to behave in diversity-supportive ways.

We saw a significant increase in participant's self-reported diversity-supporting behaviours over time, including in attitudes towards diversity, self-reported diversity supportive behaviour, feelings of psychological safety, and self-efficacy and behavioural intentions. However, except for self-efficacy and behavioural intentions, these changes were not significantly greater among participants who received training. The increase across time is likely accounted for by increased salience of diversity-related attitudes, behaviours, and other stimuli in the environment. Just drawing the attention of participants to diversity as an issue resulted in improvements in diversity-supportive attitudes and behaviours.

We did not, however, see direct changes in actual self-reported behaviour at Time 2 as a result of training. This is not surprising, given the difficulties in finding behavioural changes through experimental paradigms. However, we did find that increases in self-reported diversity behaviour at Time 2 for participants whose diversity efficacy-intentions were significantly increase by the online unconscious bias training program.

Organisations that implement unconscious bias training should focus on and measure diversity efficacy and intentions of trainees completing the program and expect to see a lag between delivery of training and behavioural improvements.

Technical report full reference list

- Ardichvili, A., Mitchell, J. A., & Jondle, D. (2009). Characteristics of ethical business cultures. *Journal of business ethics*, 85(4), 445-451.
- Ameri, M., Schur, L., Adya, M., Bentley, F. S., McKay, P., & Kruse, D. (2018). The disability employment puzzle: A field experiment on employer hiring behavior. *ILR Review*, 71(2), pp. 329-364.
- Araten-Bergman, T. (2016). Managers' hiring intentions and the actual hiring of qualified workers with disabilities. *The International Journal of Human Resource Management*, 27(14), pp. 1510-1530.
- Åslund, O., & Nordström-Skans, O. (2007). Do anonymous job application procedures level the playing field? Working Paper, IFAU - Institute for Labour Market Policy Evaluation, No. 2007:31.
- Australian Human Rights Commission. (2016). *Willing to work: National inquiry into employment discrimination against older Australians and Australians with disability*. Sydney, Australia: Australian Human Rights Commission.
- Australian Human Rights Commission. (2018). *Leading for change: A blueprint for cultural diversity and inclusive leadership revisited*. Sydney, Australia: Australian Human Rights Commission.
- Baert, S., De Visschere, S., Schoors, K., Vandenberghe, D., & Omev, E. (2016). *First depressed, then discriminated against?*. *Social Science & Medicine*, pp. 170, 247-254.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. Macmillan.
- Barta, T., Kleiner, M., & Neumann, T. (2012, April). McKinsey Quarterly: *Is there a payoff from top-team diversity?* (n.p.) McKinsey & Company. Retrieved from <https://www.mckinsey.com/business-functions/organisation/our-insights/is-there-a-payoff-from-top-team-diversity>
- Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *The American Economic Review*, 94(4), pp. 991-1013.
- Behaghel, L., Crépon, B., & Le Barbanchon, T. (2012). *Do anonymous resumes make the field more even? Evidence from a randomized field experiment*. Retrieved 26 February 2017, from <http://www.sole-jole.org/12101.pdf>.
- Bezrukova, K., Spell, C. S., Perry, J. L., & Jehn, K. A. (2016). A meta-analytical integration of over 40 years of research on diversity training evaluation. *Psychological Bulletin*, pp. 142(11), pp. 1227-1274.
- Bøg, M. & Kranendonk, E. (2011). *Labor market discrimination of minorities? Yes, but not in job offers*. Erasmus School of Economics. Retrieved 26 February 2017, from <http://mpira.ub.uni-muenchen.de/33332/>.
- Booth, A. L., Leigh, A., & Varganova, E. (2012). Does ethnic discrimination vary across minority groups? Evidence from a field experiment. *Oxford Bulletin of Economics and Statistics*, 74(4), pp. 547-573.
- Carlsson, M. (2010). Experimental Evidence of Discrimination in the Hiring of First and Second generation Immigrants. *Labour*, 24(3), 263-278.
- Carlsson, M., & Rooth, D. O. (2007). Evidence of ethnic discrimination in the Swedish labor market using experimental data. *Labour Economics*, 14(4), 716-729.
- Colic-Peisker, V., & Tilbury, F. (2006). Employment niches for recent refugees: segmented labour market in twenty-first century Australia. *Journal of Refugee Studies*, 19(2), pp. 203-229.
- Darcy, S., Taylor, T., & Green, J. (2016). 'But I can do the job': examining disability employment practice through human rights complaint cases. *Disability & Society*, 31(9), pp. 1242-1274.
- Department of Education, Employment, and Workplace Relations. (2009). *Annual report 2008-2009*. Canberra, Australia: Department of Education, Employment, and Workplace Relations. Retrieved from https://docs.education.gov.au/system/files/doc/other/annual_report_200809_part_1.pdf.
- De Plevitz, L. R. (2000). *The failure of Australian Legislation on indirect discrimination detect the systemic racism which prevents Aboriginal people from fully participating in the workforce* (PhD dissertation). Queensland University of Technology, Queensland, Australia.
- Eagly, A., Johannsen-Schmidt, M., & van Engen, M. (2003). Transformational, transactional, and laissez-faire leadership styles: A meta-analysis comparing women and men. *Psychological Bulletin*, 129 (4), pp. 569-591.
- Edin, P. & Lagerström, J. (2006). *Blind dates: quasi-experimental evidence on discrimination*. Institute for Labour Market Policy Evaluation. Retrieved 26 February 2017, from <http://www.ifau.se/globalassets/pdf/se/2006/wp06-04.pdf>.
- Ellis, K., & Sonnenfeld, J. (1994). Diverse approaches to managing diversity? *Human Resource Management*, 33, 79-109. <http://dx.doi.org/10.1002/hrm.3930330106>.
- Emerson, J. (2017, April). Don't give up on unconscious bias training - make it better. *Harvard Business Review*. Retrieved from <https://hbr.org/2017/04/dont-give-up-on-unconscious-bias-training-make-it-better>
- Forscher, P. S., Mitamura, C., Dix, E. L., Cox, W. T., & Devine, P. G. (2017). Breaking the prejudice habit: Mechanisms, timecourse, and longevity. *Journal of Experimental Social Psychology*, 72, pp. 133-146.
- Gaucher, G., Friesen, J., & Kay, A. C. (2011). Evidence That gendered wording in job advertisements exists and sustains gender inequality. *Journal of Personality and Social Psychology*, 101(1), pp. 09-128.
- Harrison, D., Price, K. H., & Bell, M. P. (1998). Beyond relational demography: time and the effects of surface- and deep-level diversity on work group cohesion. *Academy of Management Journal*, 41(1), pp. 96-107.
- Herring, C. (2009). Does diversity pay?: Race, gender, and the business case for diversity. *American Sociological Review*, 74(2), pp. 208-224.
- Hunt, V., Prince, S., Dixon-Fyle, S., & Yee, L. (2018). *Delivering through Diversity*. (n.p.) McKinsey & Company.
- Ho, C., & Alcorso, C. (2004). Migrants and employment: Challenging the success story. *Journal of Sociology*, 40(3), pp. 237-259.
- Janssens, M., & Zanoni, P. (2008, November 9). *What makes an organisation inclusive? Organisational practices favoring the relational inclusion of ethnic minorities in operative jobs*. Paper presented at the 21st International Association of Conflict Management, Chicago, Illinois, USA. Available at SSRN: <https://ssrn-com.ezp.lib.unimelb.edu.au/abstract=1298591> or doi: 10.2139/ssrn.1298591.
- Kalev, A., Dobbin, F., & Kelly, E. (2006). Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies. *American Sociological Review*, 71, 589-617. <http://dx.doi.org/10.1177/000312240607100404>.
- Kalinowski, Z. T., Steele-Johnson, D., Peyton, E. J., Leas, K. A., Steinke, J., & Bowling, N. A. (2013). A meta-analytic evaluation of diversity training outcomes. *Journal of Organisational Behavior*, 34(8), pp. 1076-1104.
- Koch, A., D'Mello, S., & Sackett, P. (2015). A meta-analysis of gender stereotypes and bias in experimental simulations of employment decision making. *Journal of Applied Psychology*, 100(1), pp. 128-161. DOI:10.1037/a0036734.
- Kochan, T., Bezrukova, K., Ely, R., Jackson, S., Joshi, A., Jehn, K., ... & Thomas, D. (2003). The effects of diversity on business performance: Report of the diversity research network. *Human Resource Management*, 42(1), pp. 3-21.
- Konrad, A. (2003). Defining the domain of workplace diversity scholarship. *Group and Organisation Management*, 28(1), pp. 4-17.
- Krause, A., Rinne, U., & Zimmermann, K. (2012). Anonymous job applications of fresh Ph.D. economists. *Economics Letters*, 117(2), pp. 441-444. doi:10.1016/j.econlet.2012.06.029.
- Jayne, M. E., & Dipboye, R. L. (2004). Leveraging diversity to improve business performance: Research findings and recommendations for organisations. *Human Resource Management*, 43(4), pp. 409-424
- Liebig, T. (2007). The labour market integration of immigrants in Australia. *OECD Social, Employment and Migration Working Papers*: OECD Publishing.

- Lowe, K. Kroeck, K., & Sivasubramaniam, N. (1996). Effectiveness correlates of transformational and transactional leadership: A meta-analytic review of the MLQ literature. *The Leadership Quarterly*, 7(3), 385–415. doi: 10.1016/S1048-9843(96)90027-2.
- Markus, A., & Dharmalingam, A. (2007). *Mapping social cohesion: The 2007 Scanlon Foundation Surveys*. Melbourne, Australia: Scanlon Foundation. Retrieved from <http://scanlonfoundation.org.au/wp-content/uploads/2016/11/2016-Mapping-Social-Cohesion-Report-FINAL-with-covers.pdf>.
- Milkman, K. L., Akinola, M., & Chugh, D. (2012). Temporal distance and discrimination: An audit study in academia. *Psychological Science*, 23(7), pp. 710–717.
- Mohr, T. R. (2014, August 25). Why women don't apply for jobs unless they're 100% qualified. *Harvard Business Review*. Retrieved from <https://hbr.org/2014/08/why-women-dont-apply-for-jobs-unless-theyre-100-qualified>
- Mollica, K.A. & Friedman, R.A. (2007). Backlash toward diversity initiatives: Examining the impact of diversity program justification, personal and group outcomes. *The International Journal of Conflict Management*, 15, pp. 77–102.
- Nishii, L. H. (2013). The benefits of climate for inclusion for gender-diverse groups. *Academy of Management Journal*, 56(6), pp. 1754–1774.
- Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social cognition*, 19(6), pp. 625–666.
- O'Leary, J., Groutis, D. and D'Almada-Remedios, R. (2017). Sydney, Australia: Diversity Council Australia. Retrieved from https://www.dca.org.au/sites/default/files/synopsis_-_cracking_the_glass-cultural_ceiling_available_to_public.pdf
- Paluck, E. L. & Green, G.P. (2009). Prejudice reduction: What works? A review of and assessment of research and practice. *Annual Review of Psychology*, 60, pp. 339–367
- Roberge, M., & van Dick, R. (2010). Recognizing the benefits of diversity: what and how does diversity increase group performance? *Human Resource Management Review*, 20(4), pp. 295–308.
- Roberson, Q. M. (2006). Disentangling the meanings of diversity and inclusion in organisations. *Group & Organisation Management*, 31(2).
- Ryan, M. K., & Haslam, S. A. (2005). The glass cliff: Evidence that women are over-represented in precarious leadership positions. *British Journal of Management*, 16(2), pp. 81–90.
- Shore, L. M., Randel, A. E., Chung, B. G., Dean, M. A., Holcombe Ehrhart, K., & Singh, G. (2011). Inclusion and diversity in work groups: a review and model for future research. *Journal of Management*, 37(4), pp. 1262–1289.
- Oreopoulos, P. (2009). *Why do skilled immigrants struggle in the labor market? A field experiment with six thousand resumes* (No. w15036). National Bureau of Economic Research.
- Pager, D., Bonikowski, B., & Western, B. (2009). Discrimination in a low-wage labor market: A field experiment. *American Sociological Review*, 74(5), pp. 777–799.
- PricewaterhouseCoopers (2011). Disability expectations: Investing in a better life, a stronger Australia. (n.p) PricewaterhouseCoopers. Retrieved from <https://www.pwc.com.au/industry/government/assets/disability-in-australia.pdf> .
- Victorian Equal Opportunity and Human Rights Commission & Deloitte (2012). *Waiter, is that inclusion in my soup? A new recipe to improve business performance*. (n.p.) Deloitte. Retrieved from <https://www2.deloitte.com/content/dam/Deloitte/au/Documents/human-capital/deloitte-au-hc-diversity-inclusion-soup-0513.pdf>.

VicHealth research component references

- Australian Bureau of Statistics, <http://www.abs.gov.au>
- Askehave, I., & Zethsen, K. K. (2014). Gendered Constructions of Leadership in Danish Job Advertisements. *Gender, Work & Organization*, 21(6), pp. 531–545. <https://doi.org/10.1111/gwao.12053>
- Azmat, G & Petrongolo, B 2014, 'Gender and the labor market: What have we learned from field and lab experiments?', *Labour Economics*, vol. 30, pp. 32–40.
- Booth, A & Leigh, A 2010, 'Do employers discriminate by gender? A field experiment in female-dominated occupations', *Economic Letters*, vol. 107, no. 2, pp. 236–8.
- Bartz, J. A., & Lydon, J. E. (2004). Close Relationships and the Working Self-Concept: Implicit and Explicit Effects of Priming Attachment on Agency and Communion. *Personality and Social Psychology Bulletin*, 30(11), pp. 1389–1401. <https://doi.org/10.1177/0146167204264245>.
- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 42(2), pp. 155–162.
- Bem, Sandra L, & Bem, D. J. (1973). Does Sex-biased Job Advertising "Aid and Abet" Sex Discrimination? *Journal of Applied Social Psychology*, 3(1), pp. 6–18. <https://doi.org/10.1111/j.1559-1816.1973.tb01290.x>
- Bohnet, I. (2016). *What Works: Gender equality by design*. Harvard University Press.
- Born, M. P., & Taris, T. W. (2010). The Impact of the Wording of Employment Advertisements on Students' Inclination to Apply for a Job. *Journal of Social Psychology*, 150(5), pp. 485–502. <https://doi.org/10.1080/00224540903365422>.
- Bosak, J., & Sczesny, S. (2008). Am I the Right Candidate? Self-Ascribed Fit of Women and Men to a Leadership Position. *Sex Roles*, 58, pp. 682–688. <https://doi.org/10.1007/s11199-007-9380-4>.
- Choi, N., & Fuqua, D. R. (2003). The Structure of the Bem Sex Role Inventory: A Summary Report of 23 Validation Studies. *Educational and Psychological Measurement*, 63(5), pp. 872–887. <https://doi.org/10.1177/0013164403258235>.
- Coelli, MB 2014, 'Occupational differences and the Australian gender wage gap', *Australian Economic Review*, vol. 47, no. 1, pp. 44–62.
- Colley, A., G. Mulhern, J. Maltby, A.M. Wood. (2009). The short form BSRI: Instrumentality, expressiveness and gender associations among a United Kingdom sample. *Personality and Individual Differences*, 46, pp. 384–387.
- Deaux, K., & L. Lewis, L. (1984). Structure of Gender Stereotypes: Interrelationships Among Components and Gender label. *Journal of Personality and Social Psychology*, 46, pp. 991–1004. <https://doi.org/10.1037/0022-3514.46.5.991>
- Diekman, A. B., & Goodfriend, W. (2006). Rolling with the Changes: A Role Congruity Perspective on Gender Norms. *Psychology of Women Quarterly*, 30(4), pp. 369–383. <https://doi.org/10.1111/j.1471-6402.2006.00312.x>
- Eagly, Alice H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3), pp. 573–598. <https://doi.org/10.1037//0033-295X.109.3.573>.
- Eagly, Alice H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, 46(4), pp. 735–754. <https://doi.org/10.1037/0022-3514.46.4.735>.
- Eagly, Alice Hendrickson, & Carli, L. L. (2007). *Through the Labyrinth: The Truth about how Women Become Leaders*. Harvard Business Press.
- Erhardt, N. L., Werbel, J. D., & Shrader, C. B. (2003). Board of Director Diversity and Firm Financial Performance. *Corporate Governance: An International Review*, 11(2), pp. 102–111. <https://doi.org/10.1111/1467-8683.00011>.

- Formanowicz, M., Bedynska, S., Cislak, A., Braun, F., & Sczesny, S. (2013). Side effects of gender-fair language: How feminine job titles influence the evaluation of female applicants. *European Journal of Social Psychology*, 43(1), 62–71. <https://doi.org/10.1002/ejsp.1924>.
- Gaucher, D., Friesen, J., & Kay, A. C. (2011). Evidence That Gendered Wording in Job Advertisements Exists and Sustains Gender Inequality. *Journal of Personality & Social Psychology*, 101(1), pp. 109–128.
- Goldin, C., Kerr, S. P., Olivetti, C., & Barth, E. (2017). The Expanding Gender Earnings Gap: Evidence from the LEHD-2000 Census. *American Economic Review*, 107(5), pp. 110–114. <https://doi.org/10.1257/aer.p20171065>.
- Gough, H. G., & Heilbrun, A. B. (1965). The Adjective Check List manual — University of Minnesota Press [Book Reference]. Retrieved September 15, 2017, from http://www.upress.umn.edu/test-division/bibliography/1960-1969/1965/gough_adjective_1965.
- Haines, E. L., Deaux, K., & Lofaro, N. (2016). The Times They Are a-Changing ... or Are They Not? A Comparison of Gender Stereotypes, 1983–2014. *Psychology of Women Quarterly*, 40(3), pp. 353–363. <https://doi.org/10.1177/0361684316634081>.
- Heilman, M. E. (2001). Description and Prescription: How Gender Stereotypes Prevent Women's Ascent Up the Organizational Ladder. *Journal of Social Issues*, 57(4), pp. 657–674. <https://doi.org/10.1111/0022-4537.00234>.
- Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, 32, pp. 113–135. <https://doi.org/10.1016/j.riob.2012.11.003>
- Heilman, M. E., & Parks-Stamm, E. J. (2007). Gender Stereotypes in the Workplace: Obstacles to Women's Career Progress. *Advances in Group Processes*, 24, pp. 47–77. [https://doi.org/10.1016/S0882-6145\(07\)24003-2](https://doi.org/10.1016/S0882-6145(07)24003-2).
- Hodel, L., Formanowicz, M., Sczesny, S., Valdov?, J., & von Stockhausen, L. (2017). Gender-Fair Language in Job Advertisements: A Cross-Linguistic and Cross-Cultural Analysis. *Journal of Cross-Cultural Psychology*, 48(3), pp. 384–401. <https://doi.org/10.1177/002202216688085>.
- Hoffman, C., & Hurst, N. (1990). Gender Stereotypes: Perception or Rationalization? *Journal of Personality and Social Psychology*, 58, pp. 197–208. <https://doi.org/10.1037/0022-3514.58.2.197>.
- Holvino, E., Ferdman, B., & Merrill-Sands, D. (2004). *Creating and sustaining diversity and inclusion in organizations: Strategies and approaches*, pp. 245–276.
- Horvath, L. K., & Sczesny, S. (2016). Reducing women's lack of fit with leadership positions? Effects of the wording of job advertisements. *European Journal of Work and Organizational Psychology*, 25(2), pp. 316–328. <https://doi.org/10.1080/1359432X.2015.1067611>.
- Ilene Wasserman, Gellagos, P., & Ferdman, B. (2008). Dancing with Resistance: Leadership Challenges in Fostering a Culture of Inclusion. In *Diversity Resistance in Organizations*. Psychology Press.
- J. Ely, R., & A. Thomas, D. (2001). Cultural Diversity at Work: The Effects of Diversity Perspectives on Work Group Processes and Outcomes. *Administrative Science Quarterly* - ADMIN SCI QUART, 46, 229–273. <https://doi.org/10.2307/2667087>.
- Jayne, M. E. A., & Dipboye, R. L. (2004). Leveraging diversity to improve business performance: Research findings and recommendations for organizations. *Human Resource Management*, 43(4), pp. 409–424. <https://doi.org/10.1002/hrm.20033>.
- Mohr, T. S. (2014, August 25). Why Women Don't Apply for Jobs Unless They're 100% Qualified. Retrieved January 23, 2018, from <https://hbr.org/2014/08/why-women-dont-apply-for-jobs-unless-theyre-100-qualified>.
- Moskos, M. (2012). 'How occupational sex segregation shapes low-skilled men's employment opportunities in Australia', *Labour & Industry: A journal of the social and economic relations of work*, vol. 22, no. 4, pp. 415–32.
- Pedriana, N., & Abraham, A. (2006). Now You See Them, Now You Don't: The Legal Field and Newspaper Desegregation of Sex-Segregated Help Wanted Ads 1965–75. *Law & Social Inquiry*, 31(4), pp. 905–938. <https://doi.org/10.1111/j.1747-4469.2006.00039.x>.
- Pelled, L. H. (1996). Demographic Diversity, Conflict, and Work Group Outcomes: An Intervening Process Theory. *Organization Science*, 7(6), pp. 615–631. <https://doi.org/10.1287/orsc.7.6.615>.
- Pelled, L. H., Eisenhardt, K. M., & Xin, K. R. (1999). Exploring the Black Box: An Analysis of Work Group Diversity, Conflict and Performance. *Administrative Science Quarterly*, 44(1), pp. 1–28. <https://doi.org/10.2307/2667029>.
- Pless, N., & Maak, T. (2004). Building an Inclusive Diversity Culture: Principles, Processes and Practice. *Journal of Business Ethics*, 54(2), pp. 129–147. <https://doi.org/10.1007/s10551-004-9465-8>.
- Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly*, 26(4), pp. 269–281. <https://doi.org/10.1111/1471-6402.t01-1-00066>
- Rosenkrantz, P., Vogel, S., Bee, H., Broverman, I., & Broverman, D. M. (1968). Sex-role stereotypes and self-concepts in college students. *Journal of Consulting and Clinical Psychology*, 32(3), pp. 287–295. <https://doi.org/10.1037/h0025909>.
- Rudman, L. A., & Glick, P. (2001). Prescriptive Gender Stereotypes and Backlash Toward Agentic Women. *Journal of Social Issues*, 57(4), 743–762. <https://doi.org/10.1111/0022-4537.00239>.
- Rudman, L. A., & Kilianski, S. E. (2000). Implicit and Explicit Attitudes Toward Female Authority. *Personality and Social Psychology Bulletin*, 26(11), pp. 1315–1328. <https://doi.org/10.1177/0146167200263001>.
- Rudman, L. A., & Phelan, J. E. (2008). Backlash effects for disconfirming gender stereotypes in organizations. *Research in Organizational Behavior*, 28, pp. 61–79. <https://doi.org/10.1016/j.riob.2008.04.003>.
- Ryan, A. M., & Kossek, E. E. (2008). Work-life policy implementation: Breaking down or creating barriers to inclusiveness? *Human Resource Management*, 47(2), pp. 295–310. <https://doi.org/10.1002/hrm.20213>.
- Schullo, S. A., & Alperson, B. L. (1984). Interpersonal phenomenology as a function of sexual orientation, sex, sentiment, and trait categories in long-term dyadic relationships. *Journal of Personality and Social Psychology*, 47(5), pp. 983–1002.
- Sczesny, S., Formanowicz, M., & Moser, F. (2016). Can Gender-Fair Language Reduce Gender Stereotyping and Discrimination? *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00025>
- Sheeran, P. & Webb, T.L. (2016, September) *The Intention-Behavior Gap*. *Social Personality Psychology Compass*, Volume 10, Issue 9, pp. 503–518.
- Stout, J. G., & Dasgupta, N. (2011). When He Doesn't Mean You: Gender-Exclusive Language as Ostracism. *Personality and Social Psychology Bulletin*, 37(6), pp. 757–769. <https://doi.org/10.1177/0146167211406434>.
- Taris, T. W., & Bok, I. A. (1998). On Gender Specificity of Person Characteristics in Personnel Advertisements: A Study Among Future Applicants. *The Journal of Psychology*, 132(6), pp. 593–610. <https://doi.org/10.1080/00223989809599292>.
- Workplace Gender Equality Agency 2016a, Australia's gender equality scorecard: Key findings from the Workplace Gender Equality Agency's 2015–16 reporting data, WGEA, p. 9, viewed 22 July 2018, www.wgea.gov.au/sites/default/files/80653_2015-16-gender-equality-scorecard.pdf.
- Vervecken, D., Hannover, B., & Wolter, I. (2013). Changing (S) expectations: How gender fair job descriptions impact children's perceptions and interest regarding traditionally male occupations. *Journal of Vocational Behavior*, 82(3), pp. 208–220. <https://doi.org/10.1016/j.jvb.2013.01.008>.

Appendix A: Hall & Wilcox CV de-identification

Appendix A contains graphs displaying the number of applicants of each gender at different stages of the hiring process for each year of data collection (1a-c) as well as the percentage of male and female applicants at each stage of recruitment for each year of data collection (2a-c). These graphs demonstrate that while female applicants comprise the majority of applicants for seasonal clerk roles at Hall & Wilcox, the overrepresentation of women does not persist across all stages of hiring, with women's overrepresentation decreasing at the hiring stage in all years.

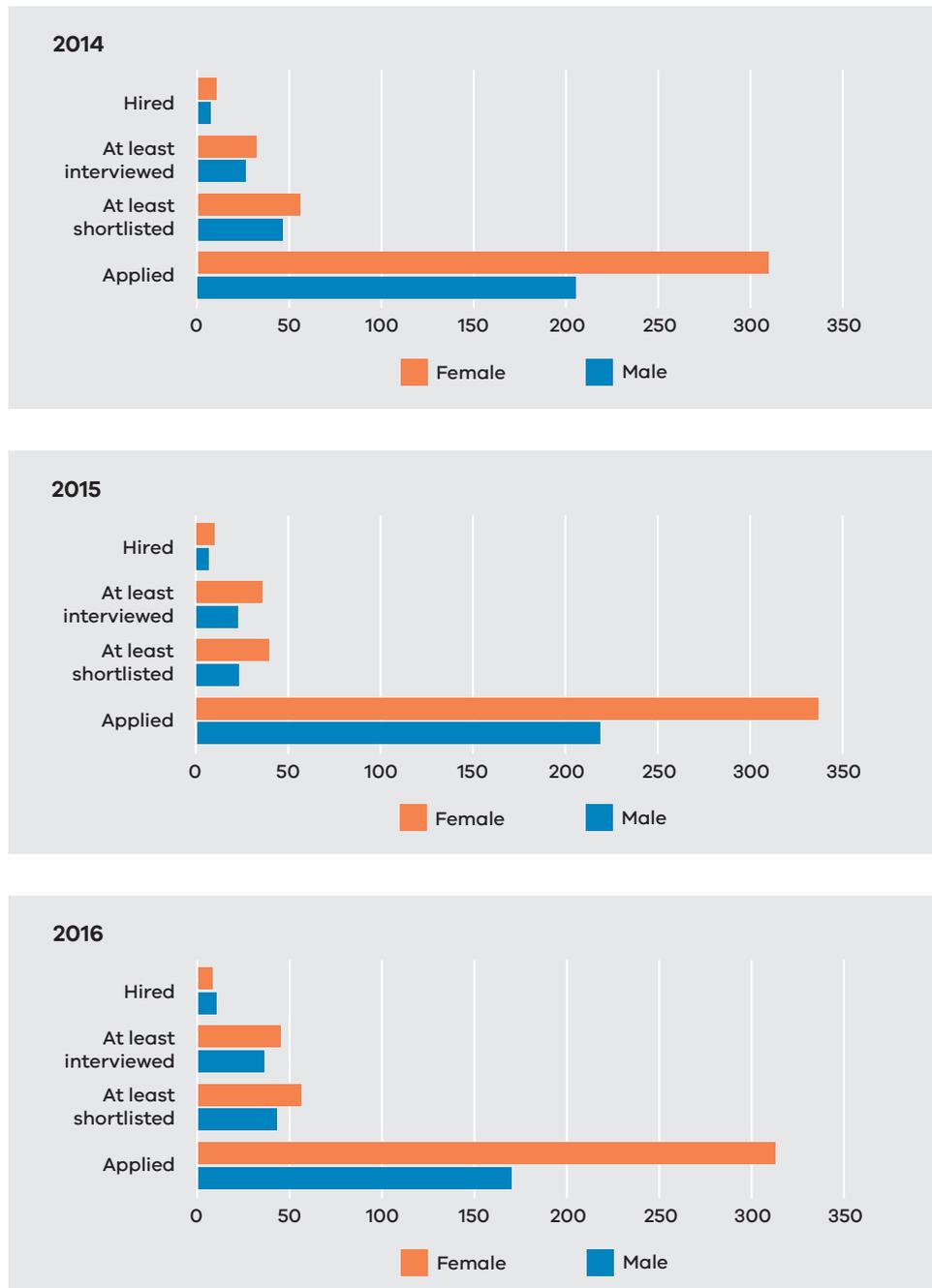


Figure 16a - c. Male and female applicant numbers at each stage of the hiring process across three years of data collection. 2014-2015 were part of the baseline condition and 2016 was part of the de-identified condition.

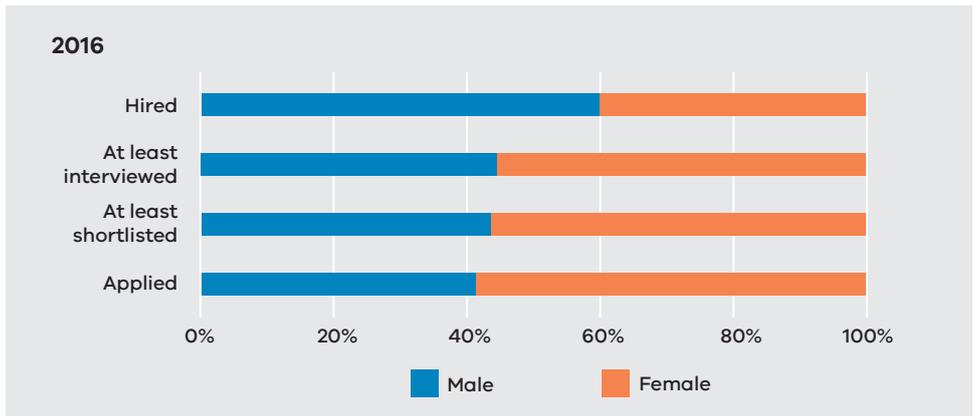
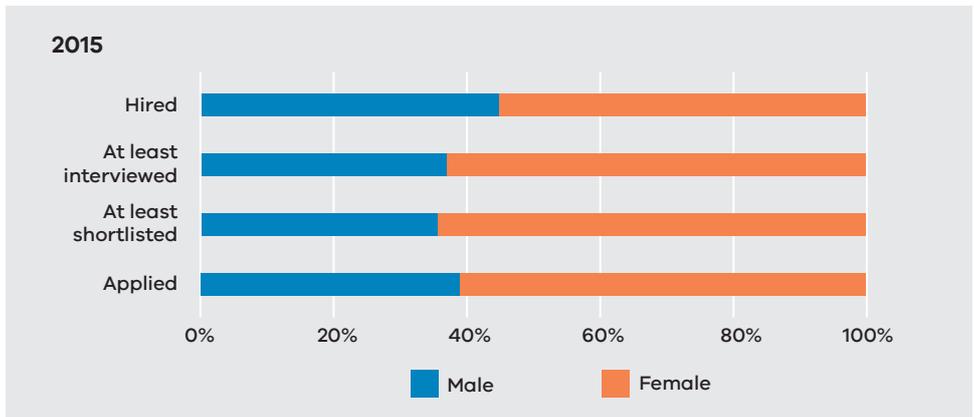
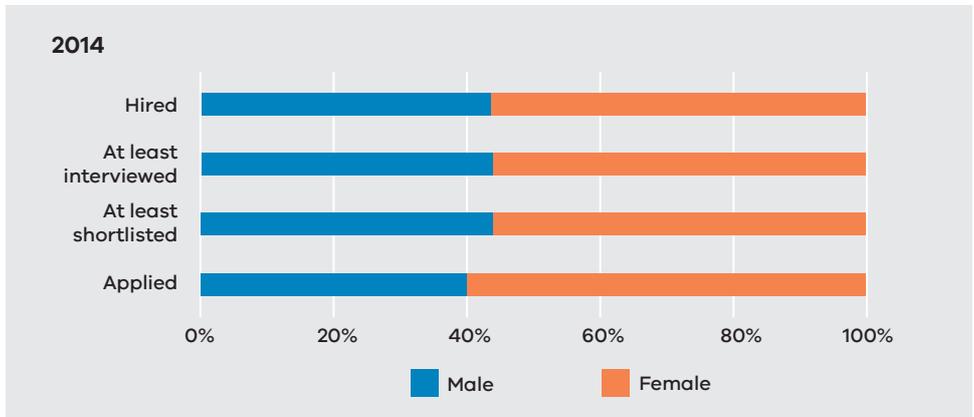


Figure 17a - c. Male and female applicant percentages at each stage of the hiring process across three years of data collection. 2014-2015 were part of the baseline condition and 2016 was part of the de-identified condition.

Appendix B: Unconscious bias training

Table 16

Unconscious Bias Training survey items and factor loadings

| REPORTED BEHAVIOUR |
|--|
| "In the past week/since completing the e-learning training modules/since completing the first survey, I have participated in the following organisational diversity efforts (please select all that apply)." |
| I have read a news article or blog, watched a television or video segment, or listened to a podcast about diversity and inclusion |
| I have attending (sic) a talk or information session about diversity and inclusion |
| I have initiated a conversation about diversity and inclusion with a colleague at work or a supervisor |
| I have initiated a conversation about diversity and inclusion with a friend or family member |
| I have reported an instance of discrimination or bullying |
| I have intervened when I witnessed discrimination or bullying |
| I changed the way I acted with others to be more inclusive |
| I challenged the biased comments of a colleague/colleagues |
| I have supported a colleague who experienced bias against him/her |
| I have identified examples of unconscious bias at work (e.g. stereotyping of a colleague) |
| My colleagues or co-workers have made critical or negative remarks about others in my workplace |
| My colleagues or co-workers have made positive remarks about others in my workplace |
| Other |

ATTITUDES

Response scale: 1 = Strongly Disagree to 5 = Strongly Agree

I am not personally concerned about a lack of diversity

People need to focus more time and energy on diversity

People make more fuss about diversity than is necessary

I consider lack of diversity at work to be a serious social problem

In the next six months I intend to contribute to and support strategies and decisions to increase workplace diversity

ORGANISATIONAL FAIRNESS

Response scale: 1 = Strongly Disagree to 5 = Strongly Agree

In your organisation or industry, people are hired because of their skills and abilities

In your organisation or industry, the most qualified candidate is the one who gets the job

In your organisation or industry, discrimination is not an issue

In your organisation or industry, diverse and underrepresented populations have a voice

In your organisation or industry, opportunities are distributed equally

PSYCHOLOGICAL SAFETY

Response scale: 1 = Very Inaccurate to 7 = Very Accurate

I am able to bring up problems and tough issues in my organisation

It is safe to take a risk in this organisation

It is easy for me to ask other members of this organisation for help

No one in this organisation would deliberately act in a way that undermines my efforts

SELF-EFFICACY AND BEHAVIOURAL INTENTIONS

"A number of workplace behaviours that relate to diversity are described below. Please rate how certain you are that you can do these behaviours on a regular basis."

Response scale: 1 = Cannot do at all to 5 = Certainly can do

Notice and question stereotypes

Advocate for diversity in the workplace

Support your organisation's diversity policies

Allocate work to employees fairly

Work in a way that includes employees equally in all discussions, decisions and tasks

Consider the impact of decisions on the performance and well-being of all employees

Try to recognise and minimise biases when analysing situations and behaviours

"Do you intend to engage in the following behaviours in the next 6 months?"

Response scale: 1 = Will not engage at all to 5 = Will definitely engage

Speak up and challenge discriminatory humour

Express appreciation for the ideas and contributions of all employees

Listen attentively and acknowledge the opinions of all employees

Share information equally with all colleagues

Facilitate or support flexible work arrangements to the extent that is possible within the organisation

Give the same opportunities for all employees to work on important projects (or support these efforts if you are not in a supervisory position).

Take steps to minimise bias in my own judgments

Notice and challenge bias in the judgments and decisions of my colleagues/team

