

Challenges to Reducing Social Bias: Predictions for a New Post Hoc Intervention

Una Tellhed

In Gunnemyr, Mattias & Jönsson, Martin L. (2023) *Post Hoc Interventions: Prospects and Problems*.
Lund: Department of Philosophy, Lund University. <https://doi.org/10.37852/oblu.184>

ISBN: 978-91-89415-60-7 (print)
978-91-89415-61-4 (digital – pdf)
978-91-89415-62-1 (digital – html)

DOI: <https://doi.org/10.37852/oblu.184.c502>



Post Hoc Interventions Prospects and Problems

Published by the Department of Philosophy, Lund University.
Edited by: Mattias Gunnemyr and Martin L. Jönsson
Cover layout by Cecilia von Arnold, Pufendorf Institute for Advanced Studies



This text is licensed under a Creative Commons Attribution-NonCommercial license. This license allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator. The license does not allow for commercial use.

(License: <http://creativecommons.org/licenses/by-nc/4.0/>)

Text © Mattias Gunnemyr and Martin Jönsson 2023. Copyright of individual chapters is maintained by the chapters' authors.

Challenges to Reducing Social Bias

Predictions for a New Post Hoc Intervention

Una Tellhed¹

Abstract. This chapter discusses a new tool for social biases interventions in work-contexts called Generalized Informed Interval Scale Update (GIIU) from a social psychological perspective. The aims are to categorize GIIU in relation to previous social bias interventions in the literature and analyze potential challenges that it may face when implemented in work contexts. Conclusions include that GIIU is a structural social bias intervention and as such, will likely meet predominately challenges that are motivational in character (the so called will not - challenge).

Aims of This Chapter

This chapter is included in an anthology that discusses a new tool for social bias interventions in work-related judgments, such as decisions for recruitment, promotion, and resource distribution. The social bias invention is called Generalized Informed Interval Scale Update (GIIU, Jönsson, 2022; Jönsson & Bergman, 2022; Jönsson & Sjö Dahl, 2017). Its method is described in the first chapter to this anthology, but I will also briefly describe it here, to enable a freestanding reading of this chapter. GI IU investigates large data materials of quantitative work-related judgments (e.g. on a 1-5 scale). An algorithm tests for mean differences related to social group categories in the material. For example, GI IU can detect if an assessor has systematically rated

¹ Una Tellhed, Senior Lecturer at the Department of Psychology, Lund University.

women's job performance as significantly lower than men's in a data set. GIIU also calculates to what degree the quantitative rating of the underrated social group could be raised, to adjust for the detected systematic difference. Assuming that performance is the same or known to be different with a certain magnitude in the two populations, GIIU can thereby correct bias "post hoc" and is therefore classified as a post hoc intervention (Jönsson, 2022; Jönsson & Bergman, 2022; Jönsson & Sjö Dahl, 2017).

This anthology takes an interdisciplinary approach to analyzing GIIU, with authors from Philosophy, Psychology, Law, Computer Science and Statistics. The current chapter analyzes GIIU from a social psychological perspective. The aim is to describe previous social bias interventions in the literature and reflect upon how GIIU fits within this literature. The main focus is to analyze common challenges that social bias interventions meet, that limits their effectiveness. Based on this previous research, I will make predictions for the type of challenges that GIIU may meet when it is implemented in organizations. I will also present some ideas for how future psychological research can investigate the implementation of GIIU. But first, I will describe what GIIU and other social bias interventions are designed to combat, namely the segregated labor market, and its roots in social bias.

The Segregated Labor Market

The Swedish labor market is strongly segregated according to sex/gender and ethnicity (Allbright, 2022; European Institute for Gender Equality, 2017; Nordic Council of Ministers, 2022; Tellhed, 2022; Wolgast & Wolgast, 2021). The lion share of research has focused on these social categories (gender and ethnicity), and therefore, so will I in this text. However, the Swedish Discrimination Act concerns the following seven social categories which are relevant for segregation in the labor market; sex, transgender identity or expression, ethnicity, religion or other belief, disability, sexual orientation and age (Government Offices of Sweden (2008:567).

The labor market is segregated both "vertically" and "horizontally" (European Commission, 2009; 2014). Vertical segregation refers to the circumstance that some social groups are disproportionally represented, relative to their population statistics, in positions of high power and status

Challenges to Reducing Social Bias

(European Commission, 2009; 2014). For example, “white”² men are overrepresented at high power positions in the labor market (Allbright, 2022, Statistics Sweden, 2022). Concerning overall employment, men and women are employed in equal numbers in Sweden, but people with an immigrant background are employed to a much lower degree as compared to people born in Sweden (Statistics Sweden, 2021).

The labor market is also horizontally segregated, which means that social groups are disproportionally represented in different types of occupations. For example, more men than women work in STEM (Science, Technology, Engineering, and Mathematics) and more women than men work in HEED (Health care, Elementary Education and Domestic (Block, et al., 2018; Nordic Council of Ministers, 2022; Tellhed, 2022)). This segregation is also partly related to status, where STEM-occupations have higher status than HEED-occupations (Croft et al. 2015; Svensson & Ulfsdotter Eriksson 2009).

The Role of Social Bias

Much research has focused on exploring why labor segregation emerges and persists. The answer is complex and outside the scope of this text, (see Tellhed, 2022 for an overview of explanations for the horizontal gender segregation), but research has shown that social bias plays one part in it (see Caleo & Heilman, 2019; Williams, 2021; Wilson, 2017; Wolgast & Wolgast, 2022 for reviews). Social bias has been defined as a

“...systematic tendency to evaluate one’s own membership group (the ingroup) or its members more favorably than a nonmembership group (the outgroup) or its members” (Hewstone et al., 2002).

Social bias is expressed as stereotypes, prejudice and discrimination, which represent cognitive, affective and behavioral aspects of bias (Hewstone et al., 2002). For example, it may include the belief that ingroup members have more meritorious qualities than outgroup members (i.e. stereotypes), having

² See Wolgast & Wolgast (2022) and Åkerlund (2022) for descriptions of how race/ethnicity is perceived in a Swedish context and the commonly perceived overlap of “whiteness” with “Swedishness”. Also, official statistics in Sweden only registers the categories “Swedish” contra “Foreign” background in demographic statistics (Statistics Sweden, 2002), which is a crude categorization as compared to race/ethnicity categorization in for example the USA.

negative attitudes or feelings towards outgroup members (i.e. prejudice), and subsequently treating ingroup members more favorable than outgroup members (i.e. discrimination and /or ingroup favoritism, Gilovich, et al., 2019; Hewstone et al., 2002). Discrimination is the behavioral aspect of the bias spectrum and is regulated in law (see chapter 5 in this anthology). Applied to a work context, examples of discriminatory behavior are selecting individuals out in the recruitment process based on their social group membership, overlooking outgroup members for promotions, or awarding outgroup members lesser rewards than ingroup members. Importantly, discrimination is more strongly related to in-group favoritism than outgroup derogation (Brewer, 1999; Greenwald & Pettigrew, 2014).

Social bias can be intentional and include conscious elements, but it can also operate without conscious access and lacking ill intent (Gawronski & Payne, 2010). Especially, when under high stress, people tend to behave automatically (without reflection or the conscious experience of intent) and may base judgments partly on stereotypes and prejudice, without taking notice (Gawronski & Payne, 2010; Pendry, & Macrae, 1994).

Social bias contributes to maintaining structural inequalities in the labor market, when its current gatekeepers select, promote and reward members of their own groups to a larger extent than outgroups. Because social bias perpetuates structural inequalities and causes career obstacles for underrepresented groups, much research has been dedicated to finding ways to mitigate social bias (e.g. Caleo & Heilman, 2019; Lai et al., 2014; Paluck & Green, 2009; Paluck et al., 2021; Williams, 2021). There are a multitude of interventions designed to reduce social bias, where some have focused specifically on increasing social diversity (i.e. reduce segregation) in the labor market (Caleo & Heilman, 2019; Lai et al., 2014; Paluck & Green, 2009; Paluck et al., 2021; Williams, 2021). I here categorize these efforts as “psychological” versus “structural” social bias interventions and will discuss their effectiveness and where GIU fits as a new addition to social bias interventions.

Psychological Social Bias Interventions

I define psychological social bias interventions as actions which aim to change individuals’ psychology in some respect, with the explicit goal to make people less biased. “Psychology” is generally described as the study of humans’ thoughts (cognition), feelings (affect), and behavior (Holt et al., 2019).

Challenges to Reducing Social Bias

Correspondingly, psychological social bias interventions aim to reduce individuals' stereotyping (cognitive bias) or prejudice (affective bias), which is then assumed to cause reductions in discrimination (behavioral bias).

Effectively reducing individuals' social bias has proven challenging. Empirical tests of psychological social bias interventions show at best moderately reduced bias post-intervention (Caleo & Heilman, 2019; Paluck & Green, 2009; Paluck et al., 2021). The longevity of the effects is rarely tested, and few studies examine behavioral outcomes (e.g. discrimination, Caleo & Heilman, 2019; Lai et al., 2017; Paluck & Green, 2009; Paluck et al., 2021). One salient concern is that some psychological social bias interventions have been found to *increase* social bias in individuals. I will describe examples of common social bias interventions and discuss their effectiveness. The categorization of the interventions is based upon previous psychological work (Caleo & Heilman, 2019; Paluck & Green, 2009; Paluck et al., 2021) and it should be noted that the concepts represent “fuzzy sets” rather than precise definitions (Mc Closkey & Glucksberg, 1978).

Diversity Training

Sociologists Frank Dobbin and Alexandra Kalev have repeatedly warned that the most common types of social bias interventions in North American work organizations, called “diversity training”, tend to have no effect, or even increase bias and segregation in organizations (e.g. Dobbin & Kalev, 2016; 2021). “Diversity training” is a fuzzily defined concept and its content varies, but one common element is informing participants about discrimination and its economic and legal consequences (Dobbin & Kalev, 2016; 2021; Paluck et al., 2021; Williams 2021). Some have proposed that shaming or threatening messages in diversity training may cause participants to react negatively to the intervention, and thereby limiting its effect (Dobbin & Kalev, 2016; 2021; Flood et al., 2021; Wiggins-Romesburg & Gibbins, 2018; Williams, 2021). Also, after being told not to discriminate certain target groups, some individuals from more privileged groups (such as white people or men) may believe that there is “reverse discrimination” in their organization, and that their ingroup is now disadvantaged relative to the groups targeted by the intervention (Dobbin & Kalev, 2016; 2021; Flood et al., 2021; Wiggins-Romesburg & Gibbins, 2018; Williams, 2021). This perception may instigate a resistance towards diversity work and increase social bias for these individuals, which I will elaborate more on later.

Post Hoc Interventions: Prospects and Problems

Although the weak results from decades of diversity training is disheartening, Joan Williams (2021) has pointed out that a new generation of diversity training programs (e.g. “habit breaking” or “bias interrupters”) show promising effects in reducing bias and increasing diversity in the workplace. These new programs teach about the psychology behind social bias, and one important difference, as compared to the previous interventions, is that it encourages staff members to come up with their own ideas for breaking bias. Enabling autonomous thinking in an intervention is more intrinsically motivating than being told not to discriminate (Devine et al., 2012; Williams, 2021).

A final thought about diversity training is that despite being taught what social bias is, and trying our best to control it, we may still discriminate against others outside of awareness (Gawronski & Payne, 2010; Pendry, & Macrae, 1994). This means that although education is important, it is not sufficient to combat segregation. There is however evidence suggesting that people who believe that they are not biased tend to discriminate more than others (Begeny et al., 2020; Régner et al., 2019). Thus, learning that social bias is common and that we all may be biased to some extent should be important. However, to complicate things further, learning that implicit (automatic) bias is common may also strengthen individuals’ bias, based on the logic that if everyone stereotypes, it is the norm (Duguid & Thomas-Hunt, 2015). This suggests that education about social bias should also try to motivate people to control their bias. Research suggests that implicit bias is controllable to some extent (Calanchini et al., 2021), but control has limitations, which I will discuss more later.

Contact Interventions

A well-researched type of psychological social bias interventions is contact interventions (Allport, 1954). Studies in this field have traditionally arranged for people from different social groups (typically ethnic groups) to meet and collaborate toward some common goal, preferably while on equal standings (Jones, & Rutland, 2018; Paluck & Green, 2009; Paluck et al., 2021). Contact interventions have been shown to moderately reduce prejudice, and the effect is related to increases in perspective taking and reduced intergroup anxiety (Aberson & Haag, 2007; Jones, & Rutland, 2018; Paluck & Green, 2009; Paluck et al., 2021).

Challenges to Reducing Social Bias

Recently, researchers have expanded contact interventions to also include “extended” or “imagined” contact. In these interventions, participants do not meet in real life, but simply read or watch material where an ingroup member is described as positively interacting with an outgroup member (Jones, & Rutland, 2018; Paluck et al., 2021). Even this minimal research design has shown prejudice reducing effects, particularly strong with children as participants, although some of the more impressive results have failed to replicate (Paluck et al., 2021).

Applied to work settings, contact interventions suggest that working collaboratively and on equal terms in diverse work teams, should reduce staff member’s social bias (Dobbin & Kalev, 2016). However, one limitation is that it demands that the organization is already diverse enough to allow for diverse work groups. Also, contact theory states that for contact to effectively reduce prejudice, group members should have equal status in the collaboration (Allport, 1954), which contrasts the current vertical segregation in the labor market that I previously described.

Social Categorization Interventions

Another common type of social bias intervention aims to alter individuals’ social categorization (Paluck & Green, 2009; Paluck et al., 2021). For example, instead of categorizing others as outgroup members (“us” versus “them”), we can try to see others as unique individuals (i.e. a lower-level categorization) or as members of a common ingroup (i.e. a higher-level categorization, e.g. “We are all employed by this company”, Paluck & Green, 2009; Paluck et al., 2021).

These types of interventions have mostly been tested in laboratory settings and tend to show effects on both implicit (indirect) and explicit measures of prejudice, although the effect sizes are typically small (Fitzgerald et al., 2019; Paluck & Green, 2009; Paluck et al., 2021).

Relatedly, interventions may also attempt to change group stereotypes by displaying counter-stereotypical examples of outgroup members (e.g. Calanchini et al., 2021; Fitzgerald et al., 2019; Lai et al., 2014; Paluck & Green, 2009; Paluck et al., 2021). However, one limitation is that when we meet an individual that counter a stereotype (such as woman with a successful career in tech), we may perceive this outgroup member as an exception, or “subtype” them, and thus preserve our stereotype intact (Kunda, & Oleson, 1995).

This implies that we need to encounter a (sufficiently) high number of outgroup members in counter-stereotypical work roles, to permanently alter

Post Hoc Interventions: Prospects and Problems

stereotypes. One interesting study showed that beginner college students associated men more than women with leadership on a computerized stereotype measure (the Implicit Association Test, Dasgupta & Asgari, 2004). However, in a follow up measure one year later, students that had encountered many women professors during their college year, now associated leadership equally strong with women as with men (Dasgupta & Asgari, 2004). This suggests that stereotypes may change when representation changes, but with the current segregation in the labor market, counter-stereotypical examples are still rare for many work roles.

Another problem with social categorization interventions has been raised in the literature of “color-blind racism” (e.g. Dovidio et al., 2015; Whitley Jr., & Webster, 2019). When majority group members claim to not “see ethnicity/race”, they may have good intentions (e.g. aspire to not be racist, Whitley Jr., & Webster, 2019). However, this strategy may have adverse effects for minority groups, which is why it is called color-blind “racism”. For example, John Dovidio and colleagues (2015) describe how color-blind approaches can create an “illusion of harmony”, where attention is distracted away from existing bias structures, while the discrimination of ethnic minority groups continues.

Social Influence Interventions

The last type of psychological social bias intervention I will discuss utilizes social influence. Research has shown that we are quite easily influenced by the opinions of our ingroup-group members, especially if they have high status, and that their comments can affect how much prejudice we express (Munger, 2017; Paluck & Green, 2009; Paluck et al., 2021; Zitek & Hebl, 2007).

Studies have tested how we are influenced by our peers with experimental design (see Paluck et al., 2009; 2021 for reviews). One interesting study used twitter bots (a software program that fakes a twitter account, Munger, 2017) to vary peer influence. The twitter bot wrote a message to white men on twitter, that had just used the n-word. The message read: “Hey man, just remember that there are real people who are hurt when you harass them with that kind of language”. When the bot portrayed a white man with many followers (high status), the users reduced their use of racist slurs. As a contrast, when the twitter bot portrayed a black man with few followers, the white men on twitter increased their racist language (Munger, 2017). Studies like this indicate that it is important that leaders and members of more privileged groups actively

Challenges to Reducing Social Bias

take part in the quest to reduce social bias in society; work that has mostly been performed by women and ethnic minorities (e.g. Caleo & Heilman, 2019).

Applying the social influence approach to working life has several limitations though. It takes courage to stand up against coworkers who express prejudice, and it may unfortunately come with a price to take part in company diversity work, due to the “resistance”, that I will discuss more later (e.g. see Caleo & Heilman, 2019; Dobbin & Kalev, 2016 for overviews).

It is also important to realize that social influence can be utilized (sometimes intentionally) to both increase as well as to decrease prejudice and inequality (Bates, 2020; Zitek & Hebl, 2007). This implies that when racist and sexist expressions are accepted in the workplace and “political correctness” is ridiculed, social bias and segregation is likely to increase.

Structural Social Bias Interventions

The research on psychological social bias interventions teaches us that it is not impossible to reduce individuals’ social bias, but that the current methods have limitations. To effectively reduce segregation in the labor market, psychological social bias interventions may be combined with “structural” social bias interventions. Structural social bias interventions are not designed to reduce individuals’ social biases per se. Instead, they can be defined as actions which change recruitment or promotion structures in some respect, with the aim to make assessors’ biases less pervasive for organizational diversity outcomes. Actions can include working towards ensuring that judgments are objective and based on merit, and/or compensating for current or historical biases that certain groups have encountered. I will exemplify two commonly described structural bias intervention methods in the literature.

Systematic Recruitment Process

Research in work and organizational psychology has developed systematic recruitment strategies which minimize the reliance on gut feeling (that is prone to bias) in decision making (Ryan & Ployhart., 2014). It is beyond the scope of this text to describe this vast field of research, but recommendations include using evidence-based test instruments (e.g. ability- and personality tests), and structured interviews in recruitment, and to analytically weigh the results from the evaluation factors into an overall judgment (Ryan & Ployhart., 2014).

Post Hoc Interventions: Prospects and Problems

When possible, anonymizing applicants further reduces the risk of social bias in judgments.

Although much research has shown that applying these recruitment methods maximizes performance outcomes in organizations, research on the “science practitioner gap” has shown that recruiters and managers are often hesitant to implement it (Highhouse, 2008; Neumann et al., 2021). Reasons include a reduced sense of autonomy over the recruitment process and limited possibilities to display one’s competence as a skillful recruiter/manager (Highhouse, 2008; Neumann et al., 2021).

The systematic recruitment methods also have limitations for the specific goal of increasing diversity, which is the focus of this chapter. Past (historical) disadvantages and discrimination that minority groups have encountered means that applicants from underprivileged groups sometimes have fewer merits, such as in education or work experience, as compared to more privileged groups (Rupp et al., 2020). Additionally, psychological phenomenon such as “stereotype threat”, means that negatively stereotyped groups may underperform on cognitive tests, due to the association of negative ability stereotypes with their ingroup (e.g. see Spencer et al., 2016; Wilson, 2017, for reviews). Previous suggestions for resolving this dilemma include weighting recruitment criteria (Rupp et al., 2020) or using quotas, to ensure representation from all target groups (Jones, et al., 2021; Roos, et al., 2020), which I will discuss next.

Affirmative Action and Quotas

In Sweden, “positive” action (i.e. affirmative action) to combat segregation in the labor market is only legal when competing candidates have equal merits, and only on the basis of sex/gender (The Equality Ombudsman, 2022a). The Swedish parliament has also voted against the EU proposal to introduce mandatory gender quotas in corporate boards (Europaportalen, 2022). Proponents for using quotas or positive/affirmative action argue that it is an efficient method to rapidly decrease segregation and that it may be necessary to compensate for historical discrimination (Jones, et al., 2021; Roos, et al., 2020). Opponents’ arguments include that it reduces corporations’ autonomy and implies decreased meritocracy (e.g. Jones, et al., 2021; Roos, et al., 2020). It has further been pointed out that quotas are also limited by the social categories it targets (e.g. sex/gender), fails to recognize intersectional power

Challenges to Reducing Social Bias

relations (e.g. how ethnicity interacts with gender) and may exclude non-binary individuals (Roos, et al., 2020).

Research has also investigated consequences for individuals recruited by quotas. One experimental study showed that when women were appointed leadership roles based on gender quotas, it lowered their perceived sense of competence, as compared to women that were recruited based on merit (Heilman et al., 1991). This problem did not occur for men that were recruited by gender quotas, possibly due to society's tendency to associate high competence with men (Storage et al., 2020). On the other hand, the same study showed that women sustained their sense of competence when recruited on quotas, if given confirmation that they have the right merits for the job (Heilman et al., 1991). This research suggests that it is important to communicate to those recruited or promoted by quotas (and to their co-workers) that they fulfill predetermined criteria for the position.

My description of psychological and structural social bias interventions has pointed to limitations with both types of interventions. I see common themes in these limitations and will next describe how they may relate to two categories of psychological functions: one motivational and one cognitive.

The “Will not” Challenge

The motivational challenge that social bias interventions face boils down to different forms of resistance towards social bias interventions and diversity work (Faludi, 1992; Wiggins-Romesburg & Githens 2018). This type of challenge affects both psychological and structural social bias interventions. Resistance towards social bias interventions has been described as one aspect of the broader term “backlash”, which is defined as resistance towards progressive social changes (Faludi, 1992; Flood et al., 2021). Much research has shown that efforts to counteract discrimination, reduce segregation and increase diversity in organizations tend to face objections, which makes progress towards increasing diversity slow (e.g. Flood et al., 2021; Lansu, et al., 2020; Wiggins-Romesburg & Githens 2018, Wilson, 2017).

Examples of identified diversity resistance strategies include denial that social bias exists or claims of reverse discrimination (that the majority group is discriminated), victim blaming or trivialization of the segregation issues, passivity and lack of engagement in anti-discrimination efforts, hidden or overt attempts to undermine anti-discrimination work, and even harassment, aggression and violence against feminists and anti-racists (Bates, 2020; Flood

Post Hoc Interventions: Prospects and Problems

et al., 2021; Jones, et al., 2021; Tildesley, et al., 2021; Wiggins-Romesburg & Githens, 2018; Wilson, 2017; Åkerlund, 2022).

Resistance to diversity work is mostly performed by individuals from normative or numerical majority groups, in organizations, such as white men (Flood et al., 2021; Williams, 2021; Åkerlund, 2022). This circumstance has been related to power motives (e.g. social dominance) and a sense of aggrieved entitlement (Flood et al., 2021; Sidanius & Pratto, 1999; Tildesley, et al., 2021; Wiggins-Romesburg & Githens 2018; Wilson, 2017). The argument is that some individuals from more privileged groups perceive that their ingroup benefits from preserving the status quo, rationalize inequalities and feel threatened when they perceive that social hierarchies are changing (Dover et al., 2016; Flood et al., 2021; Tildesley, et al., 2021; Wiggins-Romesburg & Githens 2018, Wilson, 2017).

Another strain of research describes how resistance to social bias interventions and diversity work may also stem from autonomy motivation, where some individuals reject others influencing their decision making, for example in recruitment (Dobbin & Kalev, 2016; Jones, et al., 2021; Williams, 2021). This relates to research on the science practitioner gap that I described above (Highhouse, 2008; Neumann, et al., 2021). It may also be classified as a form of power motivation (desire to control outcomes), but is more individualistic in nature, as compared to the motivation to preserve ingroup privilege.

There is a lack of research into what strategies may effectively reduce dominance-motivated resistance to social bias interventions, especially when it originates from high-power individuals (e.g. managers) in an organization (Wiggins-Romesburg & Githens 2018). If leaders of an organization passively allow for resistance to diversity work, or actively participate in it, reducing segregation is more difficult (Flood et al 2021; Lansu, et al., 2020). This implies that involving leaders and high-status individuals in diversity work is important to mitigate the resistance. Especially motivating white men with high status to participate in the diversity work should help diminish resistance from other white men in the organization, since research shows that we are mostly influenced by high-status ingroup members (e.g. Caleo & Heilman, 2019; Munger, 2017; Paluck et al., 2021). Involving employees to help find solutions to diversity problems, rather than telling them what to do, should also reduce resistance motivated by autonomy motivation (Williams, 2021). I will next turn to the other major type of limitation I see in social bias interventions.

The “Cannot” Challenge

The second psychological factor which limits the effectiveness of social bias interventions relate to cognition, and mostly concern psychological social bias interventions that aim to reduce individuals’ social bias. An important insight from social psychological research is that even when people are motivated to control their bias, they may fail to do so (Gawronski & Payne, 2010; Pendry, & Macrae, 1994). Research shows that our conscious awareness is very limited and some even argue that we have no conscious awareness of our cognitive processes, only some awareness of their output (Earl, 2014). Applied to recruitment decisions, this implies that we may have a gut feeling that we like a certain applicant better than another but limited (or no) access to what thought processes, including potential bias, that have caused these attitudes. If we are unaware of our social bias, we cannot control it.

Even when we do realize that we may have negatively stereotyped an individual, research has also shown that it is difficult to suppress stereotypes, and that attempting to do so may even increase stereotypical thinking (Macrae et al., 1994). Controlling prejudice is especially difficult when under stress, which is common in most workplaces (Pendry, & Macrae, 1994). However, actively engaging in counter-stereotypical thought is more effective than simply trying not to stereotype (Fitzgerald et al., 2019; Paluck et al., 2021) and research suggests that control attempts explain part of the reducing effect counter-stereotypical examples have on implicit bias (Calanchini et al., 2021). To change stereotypes, we also need to learn new associations, which takes practice, which is true for retraining automatized behavior in general (e.g. Calanchini et al., 2021; Gawronski & Payne, 2010).

The cognitive limitations I’ve described implies that social bias interventions that teaches participants how social bias works and instructs and motivates them to not stereotype, to mitigate their prejudice and not discriminate, may show limited success. In addition to some participants actively resisting to comply with the interventions (the “will not” challenge), reasons also include the principles of our cognitive functioning (the “cannot” challenge).

So where does GIU fit in this range of social bias interventions? What limitations do I predict for GIU and what potential may GIU have to help reduce segregation in the labor market? Also, how should future psychological research test the effectiveness of GIU, in my opinion? I will conclude with some thoughts on this matter.

Implications for GIIU

GIIU was designed as a structural intervention (Jönsson & Sjö Dahl, 2017). To remind readers, it searches through data sets with quantitative ratings of past job applications or employee performance ratings, with the aim to detect mean differences between targeted social groups. GIIU also calculates to what degree the ratings of a comparatively lower rated social group should be “corrected” assuming that the mean differences reflect assessors’ social bias. Since GIIU is not a preventative intervention (like systematic recruitment is), but detects and corrects potential biases after the fact, it has been called a “post hoc” bias intervention (Jönsson, 2022; Jönsson & Bergman, 2022; Jönsson & Sjö Dahl, 2017).

Predicted Limitations: Resistance

I predict that introducing GIIU in organizations will meet similar challenges as for other structural social bias interventions. Since GIIU is not intended to reduce assessors’ bias, but correct for it post hoc, the most relevant type of limitation is motivational resistance (the “will not” challenge).

As an example, I predict that applying GIIU in organizations will meet objections in the form of denial that detected mean social group differences reflect assessors’ social bias. That is, if GIIU for example shows that women have generally received lower ratings than men in promotion decisions, some may attribute this to women having lower competence than men, or being less career motivated than men, rather than indicating that the assessor undervalued women’s competence or merits due to bias. This assumption contradicts research which show only small gender differences or “gender similarity” in most psychological traits, including ability tests and in “agentic” (e.g. status-pursuing) career-motivation (Diekmann et al., 2016; Hyde et al., 2005, 2019; Tellhed et al., 2018; Zell et al., 2015). However, lacking evidence that an assessor has been biased by stereotypes or prejudice in their candidate ratings, other attributions are possible. For example, one may assume that a mean difference in a sample depends upon methodological limitations when GIIU was applied. Small participant samples may be skewed and not representative of population characteristics in the target categories. Further, resistance may also relate to autonomy motivation, where staff members disapprove of having their work corrected by an algorithm (Highhouse, 2008; Neumann et al., 2021).

Challenges to Reducing Social Bias

If application of GIIU will meet these types of resistance, it may imply that organizations will not use GIIU to correct for detected mean social group differences in their work-related judgments. However, organizations which hesitate to correct ratings post hoc, may still perceive the feedback from GIIU as valuable information. In Sweden, companies are obliged by law to take active measures to prevent discrimination (The Equality Ombudsman, 2022b) and GIIU may be seen as a helpful tool in this diversity work. Possibly, organizations may want to use GIIU for examining mean differences in their work-related ratings, which form the basis for their recruitment processes, promotion strategies and resource allocation. If mean differences are detected for target categories in these evaluations, it should probe for further investigation into the origins of these differences. The ratings could for example be reevaluated to ensure that they reflect differences in merits in the sample, and not assessor bias.

Future research should investigate if implementing GIIU in organizations does meet resistance and what form this potential resistance takes. It could also compare attitudes towards using GIIU as an investigative tool in organizational discrimination prevention work, versus changing the ratings post hoc in accordance with GIIU's suggestions. I suggest using both quantitative method (rating scales) and qualitative method (argument analysis) to assess attitudes and potential resistance strategies towards implementation of GIIU in organizations.

Predictors and Moderators of Attitudes

Attitudes towards GIIU, and different forms of resistance strategies, is likely to vary between staff members in organizations. Drawing on past research on resistance towards other structural interventions, that I described above, staff members from higher-status groups in the organization (such as white men) should on average display more negative attitudes towards GIIU as compared to groups with lower status (Flood et al., 2021; Williams, 2021). There should also be individual differences in attitudes within these groups. For example, attitudes could vary in relation to social dominance orientation (e.g. group based power motivation, Sidanius & Pratto, 1999; Wilson, 2017), certain personality factors such as openness and agreeableness (Akrami et al., 2009), empathy (Aberson, et al., 2007), political ideology, for example regarding the GAL (Green/Alternative/Liberal)-TAN (Traditionalist/Authoritarian/Nationalist) dimensions (Solevid et al., 2021) and autonomy motivation (Highhouse,

2008; Neumann et al., 2021; Williams, 2021). Future research could test these factors as potential predictors or moderators of attitudes towards GIIU.

Another potential moderator of attitudes towards GIIU concern who performed the evaluated past ratings; oneself, someone else, or perhaps artificial intelligence (AI), where the latter is becoming increasingly common in Sweden (The Equality Ombudsman, 2022c). I predict that having one's own past ratings assessed for suspected bias generates the most negative attitudes, since it risks exposing past discriminatory behavior one has committed. Ensuring confidentiality in GIIU applications could help reduce the risk for this type of resistance. Contrastingly, I expect the most positive attitudes if GIIU is used to correct ratings made by AI. This since using AI for decisions has been criticized for the lack of transparency into the basis for some types of AI decisions, the recent insights that also AI discriminates (The Equality Ombudsman, 2022c), and for the circumstance that AI (supposedly) has no feelings that can be hurt.

Psychological Intervention?

Lastly, although GIIU is designed as a structural social bias intervention, there might also be reason to evaluate if it can be used as a psychological social bias intervention, that is if GIIU can reduce assessor's bias. My argument is that if assessors learn that they have systematically rated a target group lower than other groups in the past, some may become motivated to reduce their social bias in future ratings, at least if they rate high on factors that relate to low resistance towards social bias interventions (Calanchini et al., 2021). However, as for other psychological social bias interventions, the effectiveness of GIIU to reduce assessors' social bias should then also depend on the limits of cognitive control that I have previously described (The "cannot" challenge).

Future research could also evaluate if GIIU: s possible potential to reduce assessors social bias may be strengthened if the GIIU output is presented in combination with education on topics such as the size of mean differences in ability in target groups and how unconscious bias operates. One could also test if adding GIIU to an existing psychological social bias intervention, such as "habit breaking training" (Devine et al., 2012) or "bias interrupters" (Williams, 2021), increases their potential to reduce individuals' social bias and organizational segregation.

If GIIU is evaluated for its potential to reduce individuals' social bias, I recommend using large enough participant samples to allow for testing of moderators of the intervention's effectiveness. Study design in psychological

Challenges to Reducing Social Bias

social bias interventions rarely include sufficiently large participant samples to do this, but it might be that the intervention has strong effects for certain individuals, but zero or even reversed effects for others (e.g. that show high resistance). When this is the case, opposite effects can cancel each other out in statistical analysis and the overall result looks weak.

Concluding Thoughts

Much research has been devoted to finding ways to reduce social bias or reducing its effect on segregation in the labor market. GIIU is a new tool that adds to this array of interventions. It is designed as a structural social bias intervention, such that it does not aim to reduce individual's biases per se but detects patterns in past work-related judgments that may have been caused by social bias, and calculates how ratings should be changed, to correct for assumed bias.

I see GIIU as a promising new tool for the quest of increasing diversity in the labor market. I predict that it will be most warmly received in the role of a potential bias detector in organizational diversity evaluations. I also predict that the function of GIIU to not only detect suspected biases in work-related decisions, but also correct for them will face resistance in organizations. However, if GIIU examines ratings made by AI, I predict that correcting suspected bias in these ratings will be more readily accepted.

To test these predictions and more, GIIU should be empirically investigated, preferably in implementation in real-world organizations. Psychological factors which may be of interest to study include attitudes towards GIIU, in-depth qualitative analysis of potential resistance in staff members, statistical testing of individual and collective factors that may relate to variation in attitudes towards GIIU, and exploration of circumstances which affects attitudes. It may also be of interest to study if GIIU is a helpful addition to current psychological social bias interventions, and may thereby contribute to reducing individuals' stereotyping, prejudice and discrimination.

GIIU is likely to meet challenges in its implementation, particularly in the form of resistance to social bias interventions. This does not mean that GIIU is defective since movement towards progressive social change will always encounter resistance. For the goal of developing a society where social group belongingness does not hinder individuals' career development, we need both psychological and structural social bias interventions. GIIU may play a role in this quest.

References

- Aberson, C. L., & Haag, S. C. (2007). Contact, perspective taking, and anxiety as predictors of stereotype endorsement, explicit attitudes, and implicit attitudes. *Group Processes and Intergroup Relations*, 10(2), 179–201. <https://doi.org/10.1177/1368430207074726>
- Akrami, N., Ekehammar, B., Bergh, R., Dahlstrand, E., & Malmsten, S. (2009). Prejudice: The person in the situation. *Journal of Research in Personality*, 43(5), 890–897. <https://doi.org/10.1016/j.jrp.2009.04.007>
- Allbright (2022). Trångsynt i toppen. [Narrow at the top]. https://static1.squarespace.com/static/5501a836e4b0472e6124f984/t/626d7a0bd81a355bc5b409c4/1651341841584/trangsynt-i-toppen_2022.pdf
- Allport G. 1954. *The nature of prejudice*. Boston, MA: Addison-Wesley.
- Bates, L. (2020). Men who hate women. The extremism nobody is talking about. Simon & Schuster.
- Begeny, C. T., Ryan, M. K., Moss-Racusin, C. A., & Ravetz, G. (2020). In some professions, women have become well represented, yet gender bias persists—Perpetuated by those who think it is not happening. *Science Advances*, 6(26), 1–10. <https://doi.org/10.1126/sciadv.aba7814>
- Block, K., Croft, A. & Schmader, T. (2018). Worth less? Why men (and women) devalue care-oriented careers. *Frontiers in Psychology*, 29(9), 1–20. <https://doi.org/10.3389/fpsyg.2018.01353>
- Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate? *Journal of Social Issues*, 55(3), 429–444. <http://dx.doi.org/10.1111/0022-4537.00126>
- Calanchini, J., Lai, C. K., & Klauer, K. C. (2021). Reducing implicit racial preferences: III A process-level examination of changes in implicit preferences. *Journal of Personality and Social Psychology*, 121(4), 796–818. <http://dx.doi.org/10.1037/pspi0000339>
- Caleo, S & Heilman, M. E. (2019). What could go wrong? Some unintended consequences of gender bias interventions. *Archives of Scientific Psychology*, 7(1), 71–80. <https://doi.org/10.1037/arc0000063>
- Croft, A., Schmader, T., & Block, K. (2015). An underexamined inequality: Cultural and psychological barriers to men’s engagement with communal roles. *Personality and Social Psychology Review*, 19(4), 343–370. <http://doi.org/10.1177/1088868314564789>

Challenges to Reducing Social Bias

- Dasgupta, N., & Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology, 40*(5), 642–658. <https://doi.org/10.1016/j.jesp.2004.02.003>
- Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. L. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology, 48*(6), 1267–1278. <https://doi.org/10.1016/j.jesp.2012.06.003>
- Diekmann, A. B., Steinberg, M., Brown, E. R., Belanger, A. L., & Clark, E. K. (2016). A goal congruity model of role entry, engagement, and exit: Understanding communal goal processes in STEM gender gaps. *Personality and Social Psychology Review, 21*(2), 142–175. <https://doi.org/10.1177/1088868316642141>
- Dobbin, F., & Kalev, A. (2016, July-August). Why diversity programs fail, and what works better. *Harvard Business Review*. <https://hbr.org/2016/07/why-diversity-programs-fail>
- Dobbin, F., & Kalev, A. (2021). The civil rights revolution at work: What went wrong. *Annual Review of Sociology, 47*, 281–303. <https://doi.org/10.1146/annurev-soc-090820-023615>
- Dover, T. L., Major, B., & Kaiser, C. R. (2016). Members of high-status groups are threatened by pro-diversity organizational messages. *Journal of Experimental Social Psychology, 62*, 58–67. <https://doi.org/10.1016/j.jesp.2015.10.006>
- Dovidio, J. F., Gaertner, S. L., & Saguy, T. (2015). Color-blindness and commonality: Included but invisible? *American Behavioral Scientist, 59*(11), 1518–1538. <https://doi.org/10.1177/0002764215580591>
- Duguid, M. M. (1), & Thomas-Hunt, M. C. (2). (2015). Condoning stereotyping? How awareness of stereotyping prevalence impacts expression of stereotypes. *Journal of Applied Psychology, 100*(2), 343–359. <https://doi.org/10.1037/a0037908>
- Earl, B. (2014). The biological function of consciousness. *Frontiers in Psychology, 5*, 1–18. <https://doi.org/10.3389/fpsyg.2014.00697>
- Europaportalen (2022, June 8). EU ett steg närmare könskvotering i bolagsstyrelser. <https://www.europaportalen.se/2022/06/eu-ett-steg-narmare-regler-om-konskvotering-i-bolagsstyrelser>
- European Commission. (2009). *Gender segregation in the labour market: Root causes, implications and policy responses in the EU*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2767/1063>

Post Hoc Interventions: Prospects and Problems

- European Commission. (2014). *A new method to understand occupational gender segregation in European labour markets*. Publications Office of the European Union. http://ec.europa.eu/justice/gender-equality/files/documents/150119_segregation_report_web_en.pdf
- European Institute for Gender Equality. (2021). *Index score for European Union for the 2021 edition*. <https://eige.europa.eu/gender-equality-index/2021/country>
- Faludi, S. (1992). *Backlash: The undeclared war against women*. London: Vintage.
- FitzGerald, C., Martin, A., Berner, D. & Hurst, S. (2019). Interventions designed to reduce implicit prejudices and implicit stereotypes in real world contexts: A systematic review. *BMC Psychology*, 7(1), 1–12. <https://doi.org/10.1186/s40359-019-0299-7>
- Flood, M., Dragiewicz, M., & Pease, B. (2021). Resistance and backlash to gender equality. *Australian Journal of Social Issues*, 56(3), 393–408. <https://doi.org/10.1002/ajs4.137>
- Gawronski, B., & Payne, K. B. (2010). *Handbook of implicit social cognition: measurement, theory, and applications*. New York: Guilford Press.
- Gilovich, T., Keltner, D., Chen, S., Nisbett, R. E (2019). Stereotyping, prejudice and discrimination. In Gilovich, T., Keltner, D., Chen, S., Nisbett, R. E (Eds.) *Social Psychology* (5th ed., pp. 359-407). New York: W. W. Norton.
- Government Offices of Sweden (2008:567). *Discrimination Act*. https://www.government.se/4a788f/contentassets/6732121a2cb54ee3b21da9c628b6bdc7/oversattning-diskrimineringslagen_eng.pdf
- Greenwald, A. G., & Pettigrew, T. F. (2014). With malice toward none and charity for some: Ingroup favoritism enables discrimination. *American Psychologist*, 69(7), 669–684. <http://dx.doi.org/10.1037/a0036056>
- Heilman, M. E., Brett, J. F., & Rivero, J. C. (1991). Skirting the competence issue: Effects of sex-based preferential selection on task choices of women and men. *Journal of Applied Psychology*, 76(1), 99–105. <https://doi.org/10.1037/0021-9010.76.1.99>
- Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup Bias. *Annual Review of Psychology*, 53(1), 575–604. <https://doi.org/10.1146/annurev.psych.53.100901.135109>
- Highhouse, S. (2008). Stubborn reliance on intuition and subjectivity in employee selection. *Industrial and Organizational Psychology-Perspectives on Science and Practice*, 1(3), 333–342. <https://doi.org/10.1111/j.1754-9434.2008.00058.x>

Challenges to Reducing Social Bias

- Holt, N., Vliek, M., Sutherland, E., Bremner, A., Passer, M., & Smith, R. E. (2019). *Psychology: the science of mind and behaviour* (Fourth edition). Maidenhead: McGraw-Hill Education
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, *60*(6), 581–592. <http://dx.doi.org/10.1037/0003-066X.60.6.581>
- Hyde, J. S., Bigler, R. S., Joel, D., Tate, C. C., & van Anders, S. M. (2019). The future of sex and gender in psychology: Five challenges to the gender binary. *American Psychologist*, *74*(2), 171–193. <https://doi.org/10.1037/amp0000307>
- Jones, S., & Rutland, A. (2018). Attitudes toward immigrants among the youth. *European Psychologist*, *23*(1), 83-92. <https://doi.org/10.1027/1016-9040/a000310>
- Jones, S.O., Taylor, S., & Yarrow, E. (2021). ‘I wanted more women in, but..’: oblique resistance to gender equality initiatives. *Work, Employment and Society*, *35*(4), 640–656. <https://doi.org/10.1177/0950017020936871>
- Jönsson, M. L. (2022). On the prerequisites for improving prejudiced ranking(s) with individual and Post Hoc interventions. *Erkenntnis*. Advanced online publication. <https://doi.org/10.1007/s10670-022-00566-2>
- Jönsson, M. L., & Bergman, J. (2022). Improving misrepresentations amid unwavering misrepresenters. *Synthese: An International Journal for Epistemology, Methodology and Philosophy of Science*, *200*. <https://doi.org/10.1007/s11229-022-03744-5>
- Jönsson, M., & Sjö Dahl, J. (2017). Increasing the veracity of implicitly biased rankings. *Episteme*, *14*(4), 499–517. <https://doi.org/10.1017/epi.2016.34>
- Kunda, Z., & Oleson, K. C. (1995). Maintaining stereotypes in the face of disconfirmation: Constructing grounds for subtyping deviants. *Journal of Personality and Social Psychology*, *68*(4), 565–579. <http://dx.doi.org/10.1037/0022-3514.68.4.565>
- Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J.-E. L., Joy-Gaba, J. A., Ho, A. K., Teachman, B. A., Wojcik, S. P., Koleva, S. P., Frazier, R. S., Heiphetz, L., Chen, E. E., Turner, R. N., Haidt, J., Kesebir, S., Hawkins, C. B., Schaefer, H. S., Rubichi, S., ... Nosek, B. A. (2014). Reducing implicit racial preferences: I A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, *143*(4), 1765–1785. <https://doi.org/10.1037/a0036260>
- Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., Calanchini, J., Xiao, Y. J., Pedram, C., Marshburn, C. K., Simon, S., Blanchar, J. C., Joy-Gaba, J. A., Conway, J., Redford, L., Klein, R. A., Roussos, G., Schellhaas, F. M. H., Burns, M., ... Nosek, B. A. (2016). Reducing implicit racial preferences:

Post Hoc Interventions: Prospects and Problems

- II Intervention effectiveness across time. *Journal of Experimental Psychology: General*, 145(8), 1001–1016. <https://doi.org/10.1037/xge0000179>
- Lansu, M., Bleijenbergh, I., & Benschop, Y. (2020). Just talking? Middle managers negotiating problem ownership in gender equality interventions. *Scandinavian Journal of Management*, 36(2), 1-9. <https://doi.org/10.1016/j.scaman.2020.101110>
- Macrae, C.N., Bodenhausen, G.V., Milne, A.B., Jetten, J. (1994). Out of mind but back in sight: stereotypes on the rebound. *Journal of Personality and Social Psychology*, 67(5): 808–17. <https://doi.org/10.1037/0022-3514.67.5.808>
- Mc Closkey, M., & Glucksberg, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory and Cognition*, 6, 462-472. <https://doi.org/10.3758/BF03197480>
- Munger, K. (2017). Tweetment effects on the tweeted: Experimentally reducing racist harassment. *Political Behavior*, 39(3), 629–649. <https://doi.org/10.1007/s11109-016-9373-5>
- Neumann, M., Niessen, A.S.M., & Meijer, R.R. (2021). Implementing evidence-based assessment and selection in organizations: A review and an agenda for future research. *Organizational Psychology Review*, 11(3), 205–239. <https://doi.org/10.1177/2041386620983419>
- Nordic Council of Ministers. (2021). Genusperspektiv på framtidens högteknologiska arbetsliv: En nordisk forskningsöversikt om utbildningsval inom STEM (science, technology, engineering and Mathematics). [A gender perspective on the high-tech work life of the future: A Nordic literature review of educational choice of STEM]. <https://www.norden.org/sv/publication/genusperspektiv-pa-framtidens-hogteknologiska-arbetsliv>
- Paluck, L. E., & Green, D. P. (2009). Prejudice reduction: What works? A review and assessment of research and practice. *Annual Review of Psychology*, 60, 339–367. <https://doi.org/10.1146/annurev.psych.60.110707.163607>
- Paluck, L. E., Porat, R., Clark, C. S., & Green, D. P. (2021). Prejudice reduction: Progress and challenges. *Annual Review of Psychology*, 72, 533–560. <https://doi.org/10.1146/annurev-psych-071620-030619>
- Pendry, L. F., & Macrae, C. M. (1994). Stereotypes and mental life: The case of the motivated but thwarted tactician. *Journal of Experimental Social Psychology*, 30(4), 303–325. <http://dx.doi.org/10.1006/jesp.1994.1015>
- Régner, I., Thinus-Blanc, C., Netter, A., Schmader, T., & Huguet, P. (2019). Committees with implicit biases promote fewer women when they do not believe gender bias exists. *Nature Human Behaviour*, 3(11), 1171–1179. <https://doi.org/10.1038/s41562-019-0686-3>

Challenges to Reducing Social Bias

- Roos, H., Mampaey, J., Huisman, J., & Luyckx, J. (2020). The failure of gender equality initiatives in academia: Exploring defensive institutional work in Flemish universities. *Gender and Society, 34*(3), 467–495. <https://doi.org/10.1177/0891243220914521>
- Rupp, D. E., Song, Q. C., & Strah, N. (2020). Addressing the so-called validity–diversity trade-off: Exploring the practicalities and legal defensibility of pareto-optimization for reducing adverse impact within personnel selection. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 13*(2), 246–271. <https://doi.org/10.1017/iop.2020.19>
- Ryan, A. M., & Ployhart, R. E. (2014). *A century of selection*. Annual Review of Psychology, *65*, 693–717. <https://doi.org/10.1146/annurev-psych-010213-115134>
- Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge: Cambridge University Press
- Solevid, M., Wängnerud, L., Djerf-Pierre, M., & Markstedt, E. (2021). Gender gaps in political attitudes revisited: the conditional influence of non-binary gender on left–right ideology and GAL-TAN opinions. *European Journal of Politics and Gender, 4*(1), 93–112. <https://doi.org/10.1332/251510820X15978604738684>
- Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype Threat. *Annual Review of Psychology, 67*, 415–437. <https://doi.org/10.1146/annurev-psych-073115-103235>
- Statistics Sweden (2002). *Statistics on persons with foreign background: Guidelines and recommendations*. <https://www.scb.se/contentassets/60768c27d88c434a8036d1fdb595bf65/mis-2002-3.pdf>
- Statistics Sweden (2021). *Stora skillnader i arbetslöshet mellan utrikes och inrikes födda* [large differences in employment between citizens born in Sweden and abroad]. <https://www.scb.se/hitta-statistik/statistik-efter-amne/arbetsmarknad/arbetskraftsundersokningar/arbetskraftsundersokningarna-aku/pong/statistiknyhet/arbetskraftsundersokningarna-aku-1a-kvartalet-2021/>
- Statistics Sweden (2022). *Women and men in Sweden - Facts and figures 2022*. https://www.scb.se/contentassets/b3ba3d3ad7a74749936c7fd2e3b4bee6/le02012021b22_x10br2201.pdf
- Storage, D., Charlesworth, T. E. S., Banaji, M. R., & Cimpian, A. (2020). Adults and children implicitly associate brilliance with men more than women. *Journal of Experimental Social Psychology, 90*, 1–14. <https://doi.org/10.1016/j.jesp.2020.104020>

Post Hoc Interventions: Prospects and Problems

- Svensson, L. G., & Ulfsdotter Eriksson, Y. (2009). *Occupational status. A sociological study on the perceptions and valuations of occupations*. (Research report no. 140). Department of Sociology, Gothenburg University, Sweden. https://gupea.ub.gu.se/bitstream/2077/19737/1/gupea_2077_19737_1.pdf
- Tellhed, U. (2022). *Val efter eget kön: Könsskillnader i utbildningsval: teori och empiri från den socialpsykologiska litteraturen*. Jämställdhetsmyndigheten. <https://jamstalldhetsmyndigheten.se/media/q0nfrnl2/bilaga-3-till-huvudrapporten-val-efter-eget-k%C3%B6n.pdf>
- Tellhed, U., Backström, M., & Björklund, F. (2018). The role of ability beliefs and agentic vs. communal career goals in adolescents' first educational choice. What explains the degree of gender-balance? *Journal of Vocational Behavior*, *104*, 1–13. <https://doi.org/10.1016/j.jvb.2017.09.008>
- The Equality Ombudsman (2022a, July 22). *Arbetslivet* [Work life]. <https://www.do.se/diskriminering/diskriminering-olika-delar-samhället/diskriminering-pa-jobbet>
- The Equality Ombudsman (2022b, September 15). *Active measures*. <https://www.do.se/choose-language/english/active-measures>
- The Equality Ombudsman (2022c). *Transparens, träning och data: myndigheters användning av AI och automatiserat beslutsfattande samt kunskap om risker för diskriminering*. <https://www.do.se/download/18.56175f8817b345aa7651be9/1646982570826/rapport-transparens-traning-och-data.pdf>
- Tildesley, R., Lombardo, E., & Verge, T. (2021). Power struggles in the implementation of gender equality policies: The politics of resistance and counter-resistance in universities. *Politics & Gender*, 1–32. <https://doi.org/10.1017/S1743923X21000167>
- Whitley, B. E., & Webster, G. D. (2019). The relationships of intergroup ideologies to ethnic prejudice: A meta-analysis. *Personality and Social Psychology Review*, *23*(3), 207–237. <https://doi.org/10.1177/1088868318761423>
- Wiggins-Romesburg, C. A., & Githens, R. P. (2018). The psychology of diversity resistance and integration. *Human Resource Development Review*, *17*(2), 179–198. <https://doi.org/10.1177/1534484318765843>
- Williams, J. C (2021). *Bias interrupted. Creating inclusion for real and for good*. Brighton, Mass.: Harvard Business Review Press
- Wilson, E. K. (2017). Why diversity fails: Social dominance theory and the entrenchment of racial inequality. *National Black Law Journal*, *26*(1), 129–153. <https://escholarship.org/uc/item/2zn704q4>

Challenges to Reducing Social Bias

- Wolgast, M., & Wolgast, S. (2021). *Vita privilegier och diskriminering: processer som vidmakthåller rasifierade ojämlikheter på arbetsmarknaden*. Länsstyrelsen i Stockholms län.
<https://www.lansstyrelsen.se/download/18.635ba3017c11a69d575fdb/1632984190659/R2021-23-Vita%20privilegier%20och%20diskriminering-webb-slutlig.pdf>
- Zell, E., Krizan, Z., & Teeter, S. R. (2015). Evaluating gender similarities and differences using metasynthesis. *American Psychologist*, 70(1), 10–20.
<http://dx.doi.org/10.1037/a0038208>
- Zitek, E. M., & Hebl, M. R. (2007). The role of social norm clarity in the influenced expression of prejudice over time. *Journal of Experimental Social Psychology*, 43(6), 867–876. <https://doi.org/10.1016/j.jesp.2006.10.010>
- Åkerlund, M. (2022). *Far right, right here: Interconnections of discourse, platforms, and users in the digital mainstream*. Akademiska Avhandlingar: Sociologiska Institutionen, Umeå Universitet.