Check for updates

Received: 17 August 2023 | Accepted: 27 August 2024

DOI: 10.1111/theo.12559

ORIGINAL ARTICLE

Reaction qualifications in the eyes of the people: An experimental-philosophical study based on US survey data

Viki Møller Lyngby Pedersen[1] | **Didde Boisen Andersen**[1,2] | **Søren Flinch Midtgaard**[1] | **Kim Mannemar Sønderskov**[1]

[1]Department of Political Science and Centre for the Experimental-Philosophical Study of Discrimination (CEPDISC), Aarhus University, Aarhus C, Denmark

[2]Vive—The Danish Center for Social Science Research, Aabyhoej, Denmark

Correspondence
Søren Flinch Midtgaard, Department of Political Science and Centre for the Experimental-Philosophical Study of Discrimination (CEPDISC), Aarhus University, Bartholins Allé 7, 8000 Aarhus C, Denmark.
Email: midtgaard@ps.au.dk

Funding information
Society of Applied Philosophy, Grant/Award Number: SAPPDS112021DBA; Danmarks Grundforskningsfond, Grant/Award Number: DNRF144

Abstract

Is it fair for employers to select candidates partly based on how the employers think customers react to the candidates' appearances, that is, based on candidates' *reaction qualifications?* Both philosophically (in the literature on wrongful discrimination) and empirically, this question has recently been getting attention. Here, we focus on a theory of unfair disadvantages emphasizing (i) whether the possession of the appearance feature in question reflects choices on the part of the candidate and (ii) whether the appearance feature in question reflects values or convictions of importance to the candidate. We scrutinize the theory by evaluating how well these factors map on to folk intuitions about wrongful discrimination using an experimental-philosophical approach and data from the United States. We do so on the presumption that coherence with folk intuitions is one desideratum for a philosophical theory and that attention should be given to potential discrepancies between philosophical theories and folk intuitions. Our central finding is that there is only general support for factor (i), not (ii), in the American population. Our study also adds to the extant empirical literature on reaction qualifications that tend to focus on the employer's side of the matter, not the candidate's as we do.

KEYWORDS

discrimination, experimental philosophy, folk intuitions, lookism, reaction qualifications

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Author(s). *Theoria* published by John Wiley & Sons Ltd on behalf of Stiftelsen Theoria.

1 | INTRODUCTION

Racism, sexism or ageism are common examples of discrimination. Yet, people can also face discrimination based on their appearance.[1] Differential and disadvantageous treatment of this kind is sometimes referred to as *lookism* (see, e.g., Mason, 2023; Monk et al., 2021; Rhode, 2010; Widdows, 2018). This phenomenon is cause for concern. An important way to express one's identity, including one's values, is through one's appearance (cf. Nuffield Council on Bioethics, 2017, p. 6). For example, some Greenlanders get facial Inuit tattoos to express and honour their Greenlandic identity. Discrimination based on these aspects of their appearance seems to deprive them of an important way of expressing themselves. Moreover, people largely form impressions of each other based on appearance, which in part explains why, for example, a person's degree of attractiveness seems to have a major impact on their job opportunities (see, e.g., Johnson et al., 2010; Bóo et al., 2013; Monk Jr. et al., 2021).

In this paper, we focus on appearance-based discrimination in the employment context. Specifically, we ask whether it is fair for employers to select candidates partly based on how the employers think customers will react to the candidate's appearance, that is, based on so-called reaction qualifications.[2] Employment based on reaction qualifications treads a fine line between meeting meritocratic norms and avoiding wrongful discrimination. Accordingly, a theory for distinguishing between legitimate and illegitimate reaction qualifications is needed. Using a survey experiment, we test a theory of unfair disadvantages emphasizing (i) whether the candidate acquired the appearance feature in question voluntarily and (ii) whether the appearance feature reflects values or convictions of importance to the candidate. These two factors are highlighted in Mason's hybrid theory of the legitimacy of reaction qualifications (Mason, 2017, pp. 56–59).

In addition to unfair disadvantage, Mason's theory also covers disrespectful mental states, disrespectful expressions and harm (including rights violations) (Mason, 2017, p. 56). In focusing on the two factors related to unfair disadvantage, we are not denying the relevance of the other elements in Mason's theory or the complexity of overall assessments of the legitimacy of reaction qualifications. Rather, our aim was to throw light on an important subset of the aspects one needs to consider when assessing the legitimacy of certain reaction qualifications.[3]

Related recent empirical studies of reaction qualifications have focused on employers' side of the matter. They find that discrimination based on reaction qualifications is considered less problematic than other forms of discrimination, hypothesizing that this is due in part to employers' limited responsibility for the adverse treatment that some groups face because of this form of discrimination: 'the employer is a collateral victim of his customers' biased preferences, and [respondents] are therefore more willing to exonerate his utilitarian behaviour' (Bunel & Tovar, 2021, p. 4701). Further, they also find that the acceptability of such discrimination is sensitive to costs. The more costly it is for an employer not to take reaction qualifications into consideration, the more acceptable it is for the employer to discriminate based on reaction qualifications (Bunel & Tovar, 2021, p. 4694, 4702). Contrary to this study, we focus on factors concerning the candidate (the voluntariness and values of the candidate).

We use data from a preregistered survey experiment conducted among a sample of the adult US population.[4] Here, we experimentally manipulated whether a fictitious job applicant with a

[1]By a person's appearance, we shall follow Mason (2023, p. 6) in meaning 'the way (or ways) in which the attributes of his or her body, and how it is adorned or modified, are (or would be) perceived visually under standard conditions by those who possess normal visual capacities together with any relevant cultural knowledge'.

[2]Important to notice, customers' reactions, and hence reaction qualifications as such, may concern features of candidates other than their appearance, for example their tone of voice. Yet, our focus here is reaction qualifications that are related to appearance.

[3]Part of our reason for focusing on unfair disadvantage is that Albertsen et al. (2023) have recently tested whether the harm- and disrespect-based theories underlying the other elements in Mason's hybrid theory reflect lay people's moral assessment of wrongful discrimination (to be clear, they also study discrimination in the employment context, but not discrimination based on reaction qualifications).

[4]The preregistration is available at https://aspredicted.org/3r6n9.pdf.

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/theo.12559 by VIVE - The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

facial tattoo (i) had gotten the tattoo voluntarily or not and (ii) whether the tattoo reflects the candidate's deeper values or not. Then, we asked respondents to evaluate whether it is legitimate for the employer to let the candidate's tattoo count as a reason against hiring the candidate. As we show in detail below, we only find support for the relevance of (i), that is the factor focusing on voluntariness. When having the tattoo was voluntary, respondents are more likely to consider it legitimate that the employer counts the tattoo as a reason against hiring the candidate than when the tattoo is acquired involuntarily. However, the increase in acceptance is small. Moreover, the results do not provide consistent evidence in favour of the importance of (ii), that is the factor focusing on the candidate's values. Interestingly, an explorative analysis of our data shows that the effects of both factors on acceptance vary across party identification. This indicates a political cleavage in people's intuitions on the matter.

We adopt an experimental-philosophical approach presuming that coherence with folk intuitions is one desideratum for a philosophical theory and that potential discrepancies between the theory and folk intuitions should be taken seriously. Motivated by the lack of consistent support for (ii), we suggest some potential revisions of the theoretical framework distinguishing between legitimate and illegitimate reaction qualifications. For example, we propose considering a reaction qualification as *less legitimate* when it reflects the candidate's deeper values (instead of *illegitimate* as Mason's theory prescribes).

The rest of our article is organized in the following way. In the next section, we present the theory of reaction qualifications on which our study is based. Moreover, we reflect on the key intuitions that underpin the theory. Our discussion revolves around a case involving a facial tattoo, which is due to Mason (2017, p. 51). On the basis of this case, we derive hypotheses regarding people's reactions to a similar case that we formulate for experimental purposes. The third section discusses the normative relevance of testing empirically whether moral principles find support in broader folk intuitions. The following sections present the research design, data and results. The final two sections discuss the results, including their philosophical implications, and give a brief conclusion.

2 | THEORY: VOLUNTARINESS AND JUDGMENTAL TASTE

Reaction qualifications are qualifications that *candidates* for jobs or positions in society may have because of how *recipients* (e.g., clients or customers) react to the candidates' features (cf. Lippert-Rasmussen, 2014, pp. 235–260; Wertheimer, 1983, p. 100). Other things being equal, a candidate is more attractive to a *selector* or an employer if the candidate has features that make recipients buy more products or services. For instance, an employer hiring salespeople for a line of beauty products or clothing may search for candidates with a beautiful appearance and a personal grooming style suggestive of fastidiousness.[5]

In a seminal article, Wertheimer asked to what extent reaction qualifications can legitimately be counted against an applicant (Wertheimer, 1983).[6] Mason (2017, 2023, Ch. 5) offers the probably most comprehensive and systematic attempt to answer the legitimacy question regarding reaction qualifications. As we do here, Mason focuses on reaction qualifications pertaining to people's appearance. According to Mason, 'a reaction qualification is *legitimate* if and only if there is no weighty moral reason not to count it, whereas it is *illegitimate* if and only if there is a weighty moral reason not to count it' (2017, p. 49). In this article, we examine one of the weighty moral reasons that Mason highlights: unfair disadvantage (Mason, 2017, p. 65). In brief, Mason proposes that counting a certain reaction qualification is illegitimate when this

[5]For a range of other examples, see Wertheimer (1983, 2006).

[6]We consistently focus on the question of the (il)legitimacy of regarding a certain appearance feature as *a* reason for rejecting a candidate (i.e. not necessarily a decisive reason). We do so to be in alignment with the theory we are testing, namely Mason's (or we are testing a crucial aspect of this theory).

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/theo.12559 by VIVE – The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

would unfairly disadvantage the candidate lacking it relative to other candidates. Whether a candidate is unfairly disadvantaged when an employer counts a reaction qualification in the hiring decision depends on two factors. The first factor concerns *voluntariness*. It matters whether the appearance feature in question is reasonably within the candidates' control (i.e. whether the feature is one that the candidates can affect in various ways). Voluntarily chosen appearance features may legitimately count against a candidate (Mason, 2017, pp. 58–59). The second factor concerns whether the feature in question reflects the agent's deeper commitments and values—or what we call the candidates' *judgmental tastes*, a term that we adopt from Cohen (2004). A judgmental taste, according to Cohen, is one that is 'informed by valuational judgement'.[7] Sometimes, appearance features (e.g., tattoos) reflect values of importance to the candidate. In such cases, it would unfairly disadvantage the candidate to count the associated reaction qualification (e.g., not having tattoos) in a hiring decision (cf. Khaitan, 2015, pp. 56–60; Mason, 2017, pp. 58–59). This applies *even if* the appearance feature is voluntarily chosen by the candidate.

Mason presents a case that falls squarely within the kind of reaction qualifications we are concerned with here:

> *Tattoo*. A man has a small but prominent facial tattoo. He values body art and thinks his tattoo is an important form of self-expression. Some people have an instinctual negative reaction to it, however. Employers in a number of fields are unwilling to hire him as a result of the disapproval his tattoo elicits (Mason, 2017, p. 51).

We draw on this tattoo case for two reasons. A tattoo is a good example of the factors that we are interested in testing: It is an appearance feature that people can have as a result of a voluntary choice, and it can reflect people's deeper values and conception of the good (Mason, 2017, p. 65). Moreover, compared with other possibly self-chosen appearance features (e.g., piercings or hair colouring), facial tattoos have a relatively permanent character and are not something that can be easily hidden or avoided while at work.[8]

The first step in an analysis of the case based on Mason's view concerns the *voluntariness* factor. Here, a central question is whether the disadvantage on the table (viz. the man's lesser employment opportunities due to his facial tattoo) is one the person could have helped having *or* could now get rid of. If either the former or the latter is true, then the disadvantage in question is the one that he himself is responsible for.[9] Bringing this criterion to bear on *Tattoo*, the case does not stipulate that the man has gotten his tattoo voluntarily. Yet barring special circumstances (and such circumstances are not mentioned in the case), the man *could* have helped having the tattoo in question. That is, he could have refrained from asking a tattoo artist to decorate his face with a prominent tattoo. Therefore, he is responsible for the disadvantage in question and cannot call for compensation. The second part of the relevant question—that is, whether he can now overcome the disadvantage (for example, have the tattoo removed)—does

[7]In contrast, so-called brute tastes are tastes that are not informed by valuational judgement. Cohen relates the distinction between judgmental tastes and brute tastes to a distinction by Dworkin between volitional and critical tastes. See Cohen, 2004: 7 and n. 11 and the references he provides there.

[8]Perhaps a tattoo can at least to some extent be covered by make-up. Yet in the case of a prominent tattoo, this is not necessarily an easy feat.

[9]Cf. Cohen (1989, p. 920): 'If he could have avoided it [a disadvantage], he has no claim to compensation, from an egalitarian point of view. If he could not have avoided it but could now overcome it, then he can ask that the effort to overcome it be subsidized, but, unless it costs more to overcome it than to compensate for it without overcoming it, he cannot expect society to compensate for his disadvantage'. We believe this to be expressive of Cohen's official stance in his 1989 article. Yet, what he in later writings refers to as the flagship statement of that article introduces some ambiguities (see Cohen, 1989, p. 923). For explanation, see Hansen and Midtgaard (2011, p. 242 n. 4).

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/theo.12559 by VIVE - The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

not arise. The fact that he could have helped having the tattoo in the first place indicates that he himself should pick up the tab for the tattoo and the disadvantages it brings.

The man has an apparently stronger fairness claim not to suffer disadvantage if he could *not* have helped having the tattoo—for example if he had acquired it as a child living in a minority culture that he has now left and distances himself from. To ascertain such a claim, we must ask whether he could now overcome the disadvantage, and he would plausibly be entitled to funds to do so. However, expecting the candidate to remove the tattoo might be considered unreasonably burdensome (compared with the burdens faced by other applicants) in terms of him having to go through a cosmetic procedure that involves both inconveniences and risks (Mason, 2017, pp. 58–59; Nuffield Council on Bioethics, 2017, Ch. 6). Accordingly, if the employer counts the tattoo against the candidate, this does seem to disadvantage him unfairly. In other words, the theory suggests that counting reaction qualifications is illegitimate when they are based on appearance features that the candidate has not incurred voluntarily.[10]

The first hypothesis——corresponding to the notion that (in)voluntariness is an important factor——is therefore:

> **H1.** When the candidate has gotten the tattoo involuntarily, people will be less inclined to agree that it is okay for the employer to let the fact that the candidate has the tattoo count as a reason against hiring the candidate (compared to when the candidate has gotten the tattoo voluntarily).

A comment regarding the wording of H_1: 'Okay' is intended to track people's assessment of the legitimacy of the employer counting the tattoo as a reaction qualification. We acknowledge that it is debatable whether it would have been better to use the mentioned term of art instead. Yet, we worried that using words such as 'legitimacy' would have made people think about what the legislation says on the matter, rather than morality.

We take support for *H1* to indicate that people consider involuntariness to be a relevant moral reason against counting the reaction qualification in disfavour of the candidate in the hiring decision.[11] On the other hand, if *H1* is rejected, this casts doubt on the notion that people consider involuntariness a relevant moral reason against counting the reaction qualification in question.

The second step in an analysis of *Tattoo* based on Mason's theory concerns the *judgmental taste* factor. Not all disadvantages due to voluntarily incurred appearance features are necessarily fair (Cohen, 2004; Segall, 2010; Stemplowska, 2009, 2011). Sometimes, agents act in certain ways because of beliefs and convictions that they cherish or that are important to them. In such cases, acting differently would be costly to them (Nuffield Council on Bioethics, 2017, p. 6). Preferences of such a judgmental taste kind should be subjected to a different standard of responsibility than should so-called brute tastes, that is preferences that people hold for no particular reason.[12]

One reason for this is that it seems problematic if people have no acceptable alternatives to incurring the costs associated with living in accordance with their values and pursuing their life

[10]The above description might be taken to suggest that 'voluntariness' is a binary phenomenon. However, we agree with authors who argue that voluntariness is a matter of degree, where the level of voluntariness involved in a decision depends on various components such as outside pressure, influence of drugs, manipulation and mistaken beliefs (see, e.g., Feinberg, 1986, ch. 20). On this view, the relevant question is whether the person acts in a manner that is *sufficiently* voluntary for the person to be held responsible.

[11]And we take this to suggest that counting an involuntarily acquired appearance feature is considered less legitimate than counting a voluntarily acquired appearance feature. Here, we deviate from the theory, which does not see legitimacy as a matter of degree (see Mason, 2017, p. 49). However, in the discussion section, we suggest that our findings may provide reason to consider legitimacy as a scale issue instead.

[12]Brute tastes are preferences that are not infused with people's convictions or judgements. Cohen mentions his own liking for Diet Coke. Such preferences are not supported by various convictions according to which this beverage is superior to other stripes of Coke or to other beverages (he just prefers Diet Coke period).

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/theo.12559 by VIVE - The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

plans (cf. Cohen, 1988, Ch. 12; Cohen, 1998; Olsaretti, 2004). Compared with the choices related to people's important plans and projects, alternative choices unrelated to these plans and projects may not seem acceptable. People are not living in ways that align with their conception of the good if they refrain from making the costly choices in question (hence, there is a sense in which making these costly choices are forced upon them) (cf. Olsaretti, 2008; Raz, 1986). All else equal, it is unfair that some people must make such costly choices (unless they would have to live contrary to their deeper commitments and values). Or at least it is when others, who hold alternative conceptions of the good, are not similarly burdened (cf. Scanlon, 1998, p. 229).

Thus, even when people choose to do something voluntarily, it is not always reasonable from a fairness perspective to let them bear the costs of their choices. According to Mason (2017, p. 65), this verdict seems especially plausible in the context of reaction qualifications where the 'cost' is 'the response a person provokes in recipients' and not a resource that is consumed in pursuit of living according to one's deeper convictions and values (e.g., not the cost of the tattoo itself).[13]

How does this argument apply to the tattoo case? First, tattoos can reflect a person's judgmental taste. For example, through their tattoos, some people may express their Greenlandic identity or their belief that it is indeed *their* body and that they can do with it as they see fit (as long as they do not use it to harm others).[14] Or, as in *Tattoo*, they may 'value body art' and think that having tattoos is an 'important form of self-expression' (Mason, 2017, p. 51). In such situations, decorating their bodies with tattoos reflects a judgmental taste. According to the theory, then it is unfair to ask the individuals to refrain from doing so (ibid., p. 65). Second, when having a tattoo reflects a person's deeper commitments and values, there is a sense in which the person is forced to make the costly decision to get a tattoo if the person wants to live in a way that reflects who the person really is. In such situations, the theory suggests that it is unreasonable to let the person incur the cost associated with being rejected for jobs because potential customers or clients will react negatively to the person's tattoo. Unless the costumers have good reasons to react negatively,[15] the theory proposes that it is unfair whether the person must choose between being rejected for the position and acting contrary to their deeper commitments and values (Mason, 2017, p. 65). That is, based on Mason's theory, it would in such circumstances be illegitimate to count the absence of tattoos as a reaction qualification.

In brief, the theory supports the view that candidates should not face disadvantages pertaining to their voluntary appearance features when these features spring from their deeper convictions and values. Disadvantages of this kind are unfair, which is a weighty moral reason against counting the reaction qualification in question. Thus, if people's intuitions on the legitimacy of reaction qualifications mirror our background theory, people will be less inclined to find it okay for an employer to reject candidates with a facial tattoo when the tattoo reflects the candidates' judgmental taste (compared to when the tattoo does *not* reflect the candidate's judgmental taste). Therefore, our second hypothesis is as follows:

> **H2.** When the tattoo forms an important part of the candidate's lifestyle and self-expression, people will be less inclined to agree that it is okay for the employer to let the fact that the candidate has the tattoo count as a reason against hiring the

[13]Moreover, it is not as if the cost of not having various employment opportunities available is connected to or somehow implied or demanded by the beliefs that underlie their decisions to get a tattoo, except, perhaps, in rare cases. For example, a person may voluntarily get a 'no future' tattoo on their forehead and believe that in fact not having a future on the job market is related to the values and beliefs that make having a tattoo important to them. In such cases, it seems less (or not at all) unfair were they to incur disadvantages on the job market. For the distinction between extrinsic and intrinsic cost, see Cohen (1989, pp. 937–939) and Mason (2017, p. 65).

[14]And assuming that doing so is permissible. For the view that this is not uncontroversial, see Cíbik (2020).

[15]Which might be considered unlikely if the tattoo is not offensive (which we stipulate that it is not in the vignettes used in the survey experiment).

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/theo.12559 by VIVE – The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

candidate (than they would be when the candidate does not consider their tattoo to be an important part of their lifestyle and self-expression).

We take support for H2 to indicate that people consider the anchoring of the appearance feature in the candidate's judgmental taste to be a relevant moral reason against counting the reaction qualification against the candidate in the hiring decision. By contrast, if H2 is rejected, this seems to challenge that the people consider the candidate's judgmental taste a relevant moral reason against counting the reaction qualification. Thus, if H2 is rejected, this suggests that there is a tension between the empirical analysis of people's intuitions and the background theory presented above. Before testing H1 and H2, we discuss the relevance of testing philosophical theories empirically.

3 | THE RELEVANCE OF FOLK INTUITIONS

What is the normative relevance of empirically testing whether the above moral principles (and by implication the philosophical intuitions underwriting them) find support in broader folk intuitions?[16] Some practitioners of experimental philosophy see the relevance of such testing as an important part of the so-called negative program—a program intended to *debunk* folk intuitions (or for that matter intuitions as such) as attractive foundations for philosophical theories. Others take a very different view and propose that experimental philosophy is the most systematic way of doing what political philosophers should be doing and that moral principles can be gleaned from such research on folk intuitions. We take a third view.[17] In contrast to the first view, we hold that studies of folk intuitions may play some positive evidentiary role regarding moral principles. Yet, in contrast to the second view, we believe that experimental philosophy is a supplementary tool, that is, not a tool that supplants more traditional philosophical methods.

According to the first view (as indicated), experimental testing may serve to reveal the, morally speaking, questionable pedigree of some of our considered judgements. For example, we may trace certain of our ('our' here meant to cover both philosophers and people more generally) moral intuitions to evolutionarily based inclinations or disinclinations. A famous example is our unwillingness to save five persons by pushing one person in front of a trolley, whereas we are much more likely to find it acceptable (perhaps even obligatory) to change tracks such that the trolley kills one person instead of five. This asymmetry in our moral intuitions might be traced to an evolutionarily based reluctance to engage in direct person-to-person physical violence (see, e.g., Greene, 2015; Greene et al., 2001; Singer, 2005).[18] Yet, it is very unclear that this reluctance provides us with compelling reasons for morally assessing the two cases differently (after all, the cases appear similar in the essential respect that they both involve the sacrifice of one person to save five) (Singer, 2005) On this view, the empirical analysis of our intuitions serves in part to *debunk* them, inviting us to look for other, sturdier foundations for making moral assessments.

The second, very different view on the relevance of empirically testing people's intuitions pertaining to philosophical theories works from the assumption that our moral intuitions are reliable. Empirical studies allow us to uncover in a systematic manner the moral principles underlying these intuitions (Kahane, 2013, pp. 428–429). As Guy Kahane (2013, p. 428) puts it:

[16]We understand intuitions as convictions regarding concrete cases that people have in part on a non-inferential basis (see Kahane, 2013, p. 422; Tersman, 2008, p. 391).

[17]For an overview of views on the relevance of x-phi for theory acceptance in political philosophy, see Midtgaard (2024).

[18]For reasons to be sceptical regarding the reliability of moral intuitions independently of experimental research see, for example, Kagan (2001, p. 53).

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/theo.12559 by VIVE – The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

> … if there are general moral principles, and our intuitions are broadly reliable, then our intuitions should be systematically responsive to the properties cited by these principles. And this means that by identifying the properties to which our intuitions systematically respond, we can identify the principles that are implicit in them.

On this view (and for that matter on the former contrasting view), there is no categorical difference between philosophers' intuitions and intuitions more broadly conceived.

We, as noted, side with the second view in taking experimental research to be relevant for moral theory acceptance. Yet in contrast to this view, we hold that there are reasons for distinguishing between philosophers' intuitions and folk intuitions. The two sorts of intuitions vary in their degree of reflectiveness. Philosophers proceeding along the lines of the method of reflective equilibrium enjoy the privilege of engaging with a morally relevant issue in an especially reflective manner. Wertheimer's original piece on reaction qualifications, for example, invented the concept and channelled further moral reflections on the issues by, for example, Mason (2017) and Lippert-Rasmussen (2014). This is a reflective process that includes consideration of a number of cases with a view to testing principles for assessing the (il)legitimacy of reaction qualifications (cf. Kauppinen, 2005, pp. 20–21, 24). Experimental settings typically allow less room for moral reflection, and the respondents presumably do not have the kind of training that philosophers have. For example, in the current study, people's reactions pertaining to a hypothetical case are examined using a survey experiment in which participants were not presented with any arguments in advance. Presumably, few of the respondents are familiar with the philosophical arguments on the (un)fairness of rejecting people based on features that are important to their values and conception of the good.

Yet, this does not imply that folk intuitions are irrelevant. Although they seem less reflective than the kind of investigation philosophers engage in, they may be sensitive to potentially relevant moral features that philosophical reflection is less sensitive to. Philosophers should at least consider such discrepancies. To mention two examples: Many philosophers find questionable the moral significance of desert towards determining what we owe each other; they are also sceptical regarding a country's right to exclude foreigners to preserve a certain conception of the nation or the communal self. People, on the other hand, think differently (Miller, 1992; Wellman, 2008, pp. 115–116). On the view we are sketching, it is incumbent upon philosophers at least to *consider* folk intuitions of this kind.

Sometimes (or perhaps often), discrepancies between philosophers' intuitions and folk intuitions do not give reason to revise theories. For example, it may be the case that folk intuitions pertaining to a country's right to exclude foreigners are driven by harmful (perhaps racist) norms. For the same reason, our view is *not* that descriptive analysis of folk intuitions should stand alone as a basis for revising theories. We place the task of formulating and developing normative theories with the philosophers. What we are saying is that knowledge of discrepancies between philosophical intuitions and folk intuitions should encourage philosophers to consider the opposing position seriously. Indeed, we find it difficult to defend the view that philosophers should not even consider opposing positions simply because these come from lay people. Taking folk intuitions into account will make the philosophical view more reasoned (even if philosophers ultimately stick to their original view). In fact, integrating an eye to folk intuitions in philosophical theorizing might be a necessary step to achieve or get closer to so-called wide reflective equilibrium (Rawls, 1995, p. 141). Favouring this third view, we conceive of folk intuitions as offering defeasible support for accepting or revising moral principles. However, the reasons in question should be seen in the context of (and weighed against) intuitions resting in part on a foundation independent of and to some extent more reflective than folk intuitions.[19]

[19]Recently, Lindauer (2020) has suggested an alternative way in which experimental philosophy may supplement standard a priori philosophical method. Lindauer appeals to the fruitfulness of normative concepts (e.g., to the kind of act they motivate people to undertake or to abstain from undertaking) and to the role of empirical research to ascertain such fruitfulness (Lindauer, 2020, esp. sec. 2).

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/theo.12559 by VIVE – The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

4 | DESIGN AND DATA

To test folk intuitions on the legitimacy of reaction qualifications, we conducted a survey experiment among a sample of the adult American population ($n = 1945$ [1856 after removing item nonresponders]). The data were collected by YouGov in March 2022 with respondents recruited from YouGov's online panel. The participants were invited based on their age, gender, geographical region of residence and education to increase representativeness of the sample.

In the survey, the respondents were presented with a hypothetical case that they were instructed to read carefully. The case described an employer considering whether to hire a candidate with a prominent, but small and nonoffensive, facial tattoo for a job involving interactions with customers (for example, a sales assistant). The employer worries that some of the customers might have negative reactions to facial tattoos and, hence, that they may be less inclined to purchase goods sold by the employer. Participants were asked whether they agree that it is okay for the employer to let the fact that the candidate has the tattoo count as a reason against hiring the candidate on a seven-point Likert scale, ranging from 1 ('strongly agree') to 7 ('strongly disagree') with 'neither disagree nor agree' in the middle. We reverse this scale in the analyses, so higher values indicate higher acceptance of letting the tattoo count as a reason against hiring the candidate, that is higher acceptance of relying on reaction qualifications.

To test whether folk intuitions about the legitimacy of relying on reaction qualifications are affected by the two factors highlighted and discussed above (involuntariness and judgmental taste), we randomly assigned the respondents to one of the three versions of the case displayed in Table 1. In the *control* condition, the candidate has gotten the tattoo voluntarily and does not consider the tattoo to be an important part of their lifestyle/self-expression. In the *involuntariness* condition, the candidate has gotten the tattoo involuntarily (but does not consider the tattoo to be an important part of their lifestyle as in the control condition). In the third condition (*judgmental taste*), the candidate does consider the tattoo to be an important part of their lifestyle but has gotten the tattoo voluntarily.

TABLE 1 Experimental conditions.

Control ($N = 620$)	**Treatment I: involuntariness ($N = 618$)**	**Treatment II: judgmental taste ($N = 618$)**
We would like you to imagine a candidate for a job involving interaction with customers (e.g., a sales assistant). The candidate has a prominent, but small and nonoffensive, facial tattoo. Apart from this, the candidate's appearance is nothing out of the ordinary. *The candidate has gotten the tattoo voluntarily. The person does not consider their tattoo as an important part of their lifestyle and self-expression.* The employer is familiar with all these facts. The employer worries that some customers might have negative reactions to facial tattoos and, hence, that they may be less inclined to make purchases from employees with such tattoos.	We would like you to imagine a candidate for a job involving interaction with customers (e.g., a sales assistant). The candidate has a prominent, but small and nonoffensive, facial tattoo. Apart from this, the candidate's appearance is nothing out of the ordinary. *The candidate has gotten the tattoo involuntarily and is not in any way responsible for getting the tattoo. The person does not consider their tattoo as an important part of their life-style and self-expression.* The employer is familiar with all these facts. The employer worries that some customers might have negative reactions to facial tattoos and, hence, that they may be less inclined to make purchases from employees with such tattoos.	We would like you to imagine a candidate for a job involving interaction with customers (e.g., a sales assistant). The candidate has a prominent, but small and nonoffensive, facial tattoo. Apart from this, the candidate's appearance is nothing out of the ordinary. *The candidate has gotten the tattoo voluntarily. The person considers their tattoo as an important part of their life-style and self-expression.* The employer is familiar with all these facts. The employer worries that some customers might have negative reactions to facial tattoos and, hence, that they may be less inclined to make purchases from employees with such tattoos.

Note: Manipulated facts emphasized in italics.

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/lheo.12559 by VIVE – The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

After reading the assigned case, the respondents were asked four questions about case details, including whether the candidate had gotten their tattoo voluntarily and whether they consider their tattoo to be an important part of their lifestyle and self-expression (see Table S1 in Supplementary Information [SI], for details). This was done to increase salience about the case and to enable an assessment of respondents' attentiveness to the case material (the assessment is reported below). These questions were followed by the outcome question on the legitimacy of relying on reaction qualifications described above. The outcome question allowed us to analyse if the variation in the case description elicits different evaluations of the legitimacy of relying on reaction qualifications. By randomly assigning participants to one of the three conditions, we ensured that nothing but the case description varied across the three groups in the aggregate.

We deliberately kept the cases short and less detailed to maximize respondents' attention and to have as little variation across the vignettes as possible. This of course leaves it open to the respondents to imagine various scenarios (e.g., how a person may have gotten a tattoo involuntarily), which admittedly carries the risk that their response might be influenced by exactly what scenario they imagine. On the other hand, presenting people with more details (for example, presenting them with a case in which a person acquired his tattoo involuntarily as a child in a minority culture) carries the risk of affecting people's responses due to factors that are not about voluntariness but other things (for example, having to do with the candidate's minority background).

As noted, the case described an employer considering whether to hire a candidate with a prominent, but small and nonoffensive tattoo. We focus on nonoffensive tattoos in that it is unclear that the theory we are testing applies to candidates' potentially offensive features. As Mason himself points out, his theory may require modifications if it is to accommodate cases where the candidates '*aim* to provoke a negative response in others' (Mason, 2017, pp. 65–66). Regarding what counts as inoffensive, we leave it up to respondents to imagine what they take to be an inoffensive tattoo. When it comes to the question whether a given tattoo is an important part of the person's self-expression, we stipulated only how the person conceives of this matter. This, of course, left it open for the respondents to question the extent to which a tattoo or an attachment to body art can be central to a person's self-conception and for the respondent to base their assessment on this disagreement. However, this was as intended because we are interested in people's view on how responsive society should be to matters that individuals and groups *claim* are important to them, that is, how tolerant society should be in this regard—at least concerning expressions that are not clearly offensive.

5 | RESULTS

In general, the respondents paid attention to their assigned case description. Across the three treatment groups, the respondents responded correctly to 2.94 of the four attentiveness questions on average (SD = 1.37; see Tables S2 and S3 in SI for details). Importantly, 69.7% of the respondents in the *involuntariness* condition responded that the candidate had gotten the tattoo involuntarily. The corresponding number was only 10.3% in the control condition, where the respondents were told that the candidate had gotten the tattoo voluntarily. This difference in proportions is statistically significant ($p < 0.001$). Similarly, significantly and substantially more respondents in the *judgmental taste* conditions reported that the tattoo is an important aspect of the candidate's lifestyle and self-expression (76.4%) than in the control condition (24.0%; $p < 0.001$). The level of attentiveness indicates that a substantial share of the respondents paid attention to the case descriptions, including the important variations regarding voluntariness and judgmental taste. This allows for a valid test of the importance of these factors for the acceptance of relying on reaction qualifications in hiring. However, we report the results of

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/theo.12559 by VIVE – The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

the experiment using both the full sample (the preregistered specification) and, as a robustness test, the results based the respondents who answered correctly on the questions regarding the two treatments. The latter results should of course be interpreted with caution given that the ability to answer correctly may correlate with other factors. We also report—again as a robustness test—the results for the full sample but with control for several covariates measured before the treatment (see SI for information about coding of the covariates).

Figure 1 and Table 2 report the results of the test of the two hypotheses. The left part of the figure shows the average level of support for relying on reaction qualifications in hiring across the three treatments (and the three different specifications). The average level of support was between 3.5 and 4 on the seven-point scale. Hence, the average level of acceptance was slightly above the midpoint of the scale (i.e. 'Neither disagree nor agree that it is okay for the employer to let the fact that the candidate has the tattoo count as a reason against hiring the candidate'; see Figure S1 in SI for the distribution of answers). The right part of the figure shows the treatment effects. Regarding the first hypothesis, there is support for the proposition that having had the tattoo involuntarily reduces the acceptance of relying on reaction qualifications. Specifically, according to the preregistered bivariate specification (shown with black circles), involuntariness significantly reduced support with 0.23 ($p = 0.04$; see Table 2). Given that the outcome is scaled from 1 to 7, the effect is small, but in the expected direction. Including covariates (the estimate shown with a square) did not change the estimate substantially, but (slightly) increases precision, as expected, given that the covariates should be uncorrelated with treatment but related to the outcome (the unrelatedness is supported by a balance test reported in Table S4 in SI, and Table 2 [Model 2] reports the associations between the covariates and the outcome). When restricting the sample to attentive respondents (coefficient with a triangle), the treatment effect not only increased slightly but also became less precise, probably due to the smaller sample. The latter coefficient is potentially biased due to selection, but it is reassuring that the result using this sample points in the same direction as the other results.

Turning to H2 regarding judgmental taste, the results did not lend support for the notion that having a facial tattoo that is considered important to the applicant's lifestyle and self-expression decreases the acceptance of relying on reaction qualifications. Across the three specifications, the average treatment effect was close to zero, in the wrong direction with regard to the hypothesis and statistically insignificant. H2 was therefore rejected.

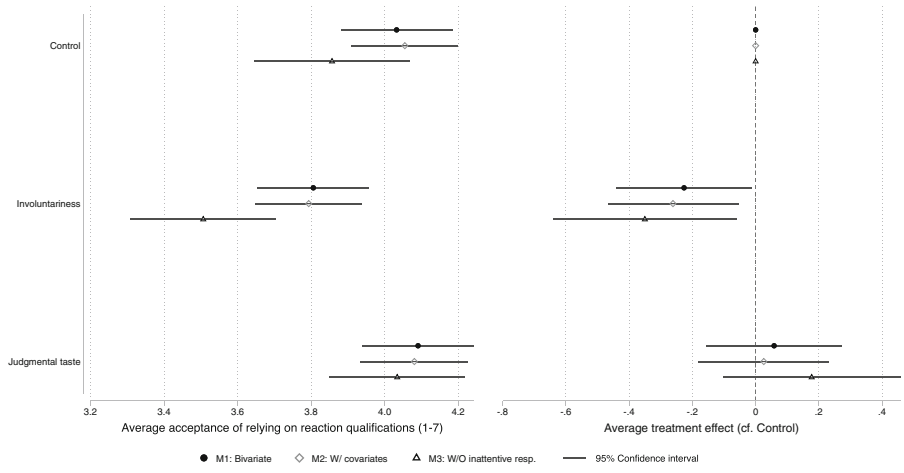


FIGURE 1 Average acceptance of relying on reaction qualifications and average treatment effects. $N = 1856$ (Models 1–2)/1162 (Model 3). See Table 2 for the full results and the list of covariates in Model 2. All covariates are held at their observed values when calculating the predicted level of acceptance based on Model 2.

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/lheo.12559 by VIVE – The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Full results behind Figure 1.

	(1)	(2)	(3)
Treatment (ref.: control)			
Involuntariness	−0.23* (−2.07)	−0.26* (−2.49)	−0.35* (−2.37)
Judgmental taste	0.06 (0.53)	0.03 (0.24)	0.18 (1.24)
Gender (ref.: female)			
Male		0.63*** (7.00)	
Age (ref.: 18–24)			
25–34		−0.09 (−0.51)	
35–44		0.04 (0.23)	
45–54		−0.10 (−0.54)	
≥55		0.11 (0.65)	
Education (ref.: highschool or below)			
Some college		−0.25* (−2.14)	
4-year college		−0.04 (−0.32)	
Post-grad		−0.13 (−0.86)	
Race (ref.: White)			
Black		0.19 (1.40)	
Hispanic		−0.01 (−0.06)	
Other		0.15 (0.99)	
Family income (ref.: < 30 K)			
30–60 K		−0.07 (−0.52)	
60–100 K		0.11 (0.77)	
100–150 K		−0.01 (−0.03)	
150–350 K		0.37* (2.01)	
≥350 K		0.81* (2.28)	
Prefer not to say		0.06 (0.39)	
Region (ref.: Northeast)			
Midwest		0.02 (0.15)	
South		0.03 (0.22)	
West		0.02 (0.16)	
Employer (ref.: no)			
Employer		0.83*** (6.32)	
Party identification (ref.: Democrat)			
Independent		0.33** (3.01)	
Republican		0.89*** (7.54)	
Other/DK/NA		0.09 (0.56)	
Have tattoo (ref.: yes)			
No		0.13 (1.30)	
Prefer not to say		0.09 (0.33)	
Constant	4.03*** (52.12)	3.21*** (14.59)	3.86*** (35.68)
Observations	1856	1856	1162

Note: *t* statistics in parentheses. See SI for details about coding of covariates.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/theo.12559 by VIVE – The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

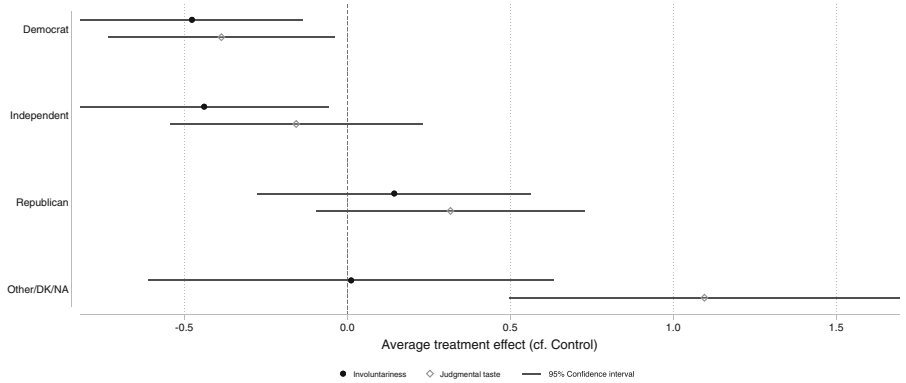


FIGURE 2 Marginal average treatment effects across political party orientations. $N = 1856$. Results are based on a model with the same specification as Model 2 in Table 2, but including an interaction term between treatment status and party identification. See Table S12 in SI for the full results. *DK* signify 'don't know' and *NA* signify 'prefer not to say'.

A relevant concern here is whether our results were a consequence of a low-powered design. However, this does not seem to be the case. A post hoc power analysis (Bloom, 1995) based on the two estimates from Model 1 suggests that our design had 80% power to detect a treatment effect of around 0.31 for both treatments, which corresponds to a Cohen' *D* of 0.16. This implies that our design had reasonable power to detect fairly small effects.

A follow-up question to the results just presented is whether the treatment effects vary across groups. Is the involuntariness aspect perhaps nonconsequential for certain groups or could the judgmental tastes aspect reduce the acceptance of reaction qualifications for others? Model 2 in Table 2 reveals that different groups indeed have different views of the legitimacy of relying on reaction qualifications in general. Males (compared with females), wealthier, employers as well as independents and in particular republicans were more accepting of relying on reaction qualifications. However, this does not tell us if the treatment effects varied across groups. To investigate this question, we interacted each of the covariates included in Model 2 in Table 2 with the treatment indicator (one covariate at the time in separate models, but with control for the other covariates).

Generally, there was not much evidence for heterogeneous treatment effects across groups (the interaction terms were not statistically significant at conventional levels; see Tables S5–S13 in SI)—with the notable exception of political party identification.[20] The analysis suggested that the treatment effects vary across party identification. Figure 2 depicts this and shows that the negative treatment effect of the *involuntariness* treatment exclusively was found among voters identifying as a democrat or an independent whereas it was insignificant (and marginally positive) among republicans and respondents identifying with other parties, answering 'don't know' or 'do not wish to say'. The analysis of heterogenous effects for the judgmental taste treatment suggests even larger effect heterogeneity as the hypothesis about a negative treatment effect actually found support among democrats and (to a lesser extend) independents. Among republicans and especially respondents identifying with other parties, answering 'don't know' or 'do not wish to say', the treatment effect was positive, which explain the overall null effects. These results should of course be interpreted with caution as we did not preregister a hypothesis regarding heterogenous effects (we only listed such analyses as explorative). The statistical power of the design to detect heterogenous effects is also limited, and political party identification was not randomly assigned, so other factors than ideology may explain the heterogeneity.

[20]There were also indications of a weak and hard-to-interpret interaction with income (see Table S9 in the SI).

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/1467-9914.12559 by VIVE - The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Nevertheless, the analysis suggested that political ideology could play an important role in relation to folk intuitions about reaction qualifications—not least in relation to the importance of voluntariness and judgmental taste. Preferably, future studies could explore this further.[21]

6 | DISCUSSION OF THE FINDINGS AND THEORETICAL IMPLICATIONS

Our first main finding is that, in the case of facial tattoos, Americans' reactions to the legitimacy of counting the relevant reaction qualification are in part influenced by considerations of voluntariness. If the candidate got the tattoo involuntarily, Americans are generally less likely to consider it legitimate for the employer to reject the candidate. In this way, the study provides support for the first factor emphasized in Mason's theory. However, as emphasized in the previous section, the decrease in acceptance if the candidate got the tattoo involuntarily is small—especially when considering the significant role that the factor plays in the theory. According to the theory, it is illegitimate to count having a tattoo as a reaction disqualification if the tattoo was incurred involuntarily. Thus, although there is support for the idea that (in)voluntariness plays a morally relevant role, it seems that the decrease in acceptance that we find does not match the theoretical importance of this factor.

It should be noted, however, that we test people's intuitions pertaining to a specific case involving appearance, specifically a facial tattoo, and we cannot exclude that people may react more strongly to counting reaction qualifications in other cases (an issue we return to below). Moreover, our test is based on American respondents only, and the acceptance of reaction qualifications based on appearance features obtained involuntarily may be lower in other cultures. The heterogenous reactions across political ideology suggest that this could be the case.[22]

Our second main finding is that the average American's intuitions pertaining to the legitimacy of counting the relevant reaction qualification do not depend on information about whether the tattoo reflects the applicant's values or judgements. Does rejecting H_2 imply that we should reject Mason's view on the legitimacy of reaction qualifications? It is a clear implication of the view that reaction qualifications based on appearance features that reflect a person's 'reasonable moral doctrine or conception of the good' are illegitimate (Mason, 2017, p. 58), and it is equally clear that the theory in question considers a facial tattoo to be part of a doctrine of this kind (Mason, 2017, pp. 59 n. 16, 65). Typically, if a theory has implications that conflict with our philosophical considered judgements, this gives us reason to reject or at least adjust the theory. Furthermore, if a theory has implications that conflict with folk intuition, then this might warrant adjustments.

Yet, again, one may doubt whether our findings pertaining to the specific tattoo case reflect 'folk intuitions' regarding the legitimacy of reaction qualifications more generally. The case of facial tattoos may differ from other traits in relevant respects. For example, even though we stipulate that the tattoo is nonoffensive, people may associate tattoos with negative values, which might affect their responses.[23] Specifically, people may be less sensitive to the judgmental taste manipulation in the tattoo context than they would have been in other contexts involving other appearance features. However, it is far from given that tattoos differ from other appearance features in this respect. For example, one might have the exact same worry when it comes

[21]In view of these heterogeneous treatment effects, it might be questioned whether we can reject H2 based on our findings. H2 concerns the *average* treatment effect, which is unaffected by the judgmental taste manipulation. We did not preregister a hypothesis concerning *heterogeneous* treatment effects.

[22]On the other hand, a similar study conducted among a representative sample of the Danish population shows congruent results (Andersen et al., 2023), which speaks against this concern.

[23]It is not the case that respondents generally agree that it is okay for the employer to count the tattoo against the candidate in the hiring decision (i.e. we do not find any ceiling effects). If people generally associate tattoos with something negative, this is not reflected in them, generally finding it okay to reject the tattooed candidate.

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/theo.12559 by VIVE - The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

to religious head coverings, which people might also associate with things (e.g., oppression of women) that they object to. Our study does not examine whether the effect of the judgmental taste manipulation is context dependent. Follow-up studies with varying appearance features would illuminate this.

Our results regarding judgmental taste may also be sensitive to the population under study. Recall that the explorative analysis of heterogeneous effects for the judgmental taste treatment suggests support for H2 among democrats and independents, whereas the effects among republicans and respondents identifying with other parties, answering 'don't know' or 'do not wish to say' point in the wrong direction relative to H2. It is possible that these differences can be explained by variation in the importance (or lack of such) that these groups attach to identity issues. In any case, the explorative findings cast doubt on the existence of a uniform 'folk intuition' regarding the relevance of the judgmental taste factor. However, for the reasons highlighted above, we should be cautious about placing too much weight on these specific findings regarding political identification.

Suppose that it *is* sensible to conclude that arguments about judgmental taste do not find support in folk intuitions. What kind of adjustments to the theory could this reasonably give rise to? One way of adjusting the theory in the light of the finding that judgmental tastes play less of a role in folk intuitions than they do in philosophers' intuitions is the following. Mason's theory proposes that a reaction qualification is *illegitimate* in the case that the reactions in question concern appearance features of the candidate that are associated with their judgements. Instead, the revised theory holds that counting reaction qualifications in the indicated circumstances is *less legitimate* than it would otherwise be. In this way, we soften somewhat the judgmental taste factor in our assessments of the legitimacy of counting reaction qualifications, and we do so in part in response to folk intuitions, but without abandoning our philosophical intuitions that take a more favourable view on the moral salience of the factor in question. So, whereas in some cases we should stick uncompromisingly to our philosophical intuitions, in other cases (including the one under consideration), we might modify our intuitions to some extent in light of folk intuitions.

For example, we may have very strong intuitions to the effect that in certain cases, sex *is* a relevant reaction qualification, including when we hire gynaecologists (because some women patients prefer female gynaecologists), and should perhaps stand firm even if folk intuitions (surprisingly) suggest otherwise (Mason, 2017, p. 50; Yuracko, 2006, pp. 867–868). In other cases, such as the present one, we may modify somewhat our intuitions partly because our intuitions are less firm: We stand in need of more guidance (cf. Rawls, 1999, p. 18), and some of that guidance might come from folk intuitions.

An alternative to adjusting the theory is to refine our philosophical analysis. Instead of taking folk intuitions regarding the relevance of judgmental tastes to affect the main question of the *(il)legitimacy* of reaction qualifications, we could understand them to affect our all-things-considered judgements concerning whether to count reaction qualifications. As Mason argues, reaction qualifications might in some cases be illegitimate, but justifiable to count all things considered (Mason, 2017, p. 67). That is, although there are weighty moral reasons against counting a certain reaction qualification, there might be countervailing reasons of a kind that can outweigh the (weighty) moral reasons in question. Specifically, although counting recipients' negative reactions to certain specific appearance features may be illegitimate, it might be justifiable to count them, all things considered. For example, it might be so if not counting these reactions would plausibly mean that the firm in question would go out of business. Accordingly, while we may have compelling philosophical reasons for considering certain appearance-related reaction qualifications *illegitimate* in part because they are reactions to appearance features that are bound up with people's judgmental tastes, these weighty reasons can be outweighed by other weighty reasons. Among these other reasons may be that, according to folk intuitions, the fact that certain appearance features relate to people's

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/theo.12559 by VIVE - The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

judgmental tastes does not in itself constitute a reason for not counting the reaction qualifications in question.

Consider a case in which the employer must operate in an area in which customers entertain very negative views about body art, especially facial tattoos. Assume also that a given candidate for a job cherishes body art, as in one of the cases we considered above. Presume finally that if the employer does not regard the absence of a facial tattoo as a reaction qualification, the employer's profit would drop considerably, but she would not go out of business. In this case, initially only consulting our philosophical intuitions, we may plausibly think that it is both *illegitimate* and *unjustifiable all things considered* to count the reaction qualification in question. After all, from this perspective, it is important that people do not experience disadvantages on the basis of appearance features that they cherish, and the employer, let us presume, can maintain her business (although with less profit than if the reaction qualification were counted).

Yet, taking into consideration folk intuitions may serve to tip the balance so that counting the reaction qualification in question remains illegitimate *but* now seems justifiable all things considered. Specifically, people find the acceptability of discrimination based on reaction qualifications sensitive to the costs associated with hiring a candidate who lacks the qualification in question (Bunel & Tovar, 2021, p. 4701), but insensitive to whether the relevant feature bears on the candidate's important values. These folk intuitions may count as relevant considerations in an all-things-considered assessment, which suffices to make counting the reaction qualification justifiable all things considered.

To summarize, we may take the findings to either affect the question regarding the legitimacy of counting reaction qualifications or to inform the question of whether counting an illegitimate reaction qualification is all things considered justifiable. In both cases, the findings affect our moral assessment of reaction qualifications in a nontrivial manner and in accordance with our methodological conception as sketched above.

In the above analysis, the employer's interest in staying in business is mentioned as a potentially weighty reason in favour of counting a given reaction qualification. This view seems consistent with the theory's focus on voluntariness. The candidate having the tattoo and the customers' negative reactions to it may reasonably be seen as something that is beyond the control of the employer. Accordingly, given the theory, it seems unfair, all things equal, if the employer were to incur the costs associated with the candidate's avoidable (although perhaps not reasonably avoidable) tattoo (see also, Bunel & Tovar, 2021, p. 4694).

On the other hand, if there is no law or policy that requires the employer not to discriminate against candidates based on appearance features that reflect their membership of a certain protected group (as, e.g., skin colour does), and the employer chooses to hire the candidate in question, this can be ascribed to the employer's voluntary choices.[24] Thus, in such circumstances the disadvantages the employer might face if they hire a candidate with facial tattoos do not seem to be unfair.

However, just as having a facial tattoo may be grounded in the candidate's judgmental tastes, the choice to hire a candidate with a facial tattoo may be grounded in the employer's judgmental tastes. The employer may, for example, value diversity and find it wrongful to reject candidates based on other people's prejudices. In that case, the disadvantages associated with hiring the tattooed candidate are avoidable to the employer but not *reasonably* so (cf. Section II). Accordingly, on the unfair disadvantages view, we can easily end up in a situation where it would both be unfair that the candidate is disadvantaged if the reaction qualification is counted *and* unfair that the employer is disadvantaged when the reaction qualification is not counted (and in this case, appealing to folk intuitions would be to no avail in that this would suggest reducing somewhat the importance of the judgmental tastes of both the

[24]Option luck, as defined by Dworkin (2000, p. 73), 'is a matter of how deliberate and calculated gambles turn out—whether someone gains or loses through accepting an isolated risk he or she should have anticipated and might have declined'.

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/theo.12559 by VIVE - The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

candidate and the employer). One possible response to this dilemma is to introduce compensation to the unfairly disadvantaged party. We might think that if the reaction qualification is counted, the candidate should be compensated, and if it is not counted, the employer should receive some kind of compensation.

How would this analysis look if the candidate were disadvantaged because of customers' or clients' prejudices towards members of certain socially salient groups (e.g., if clients are racists or sexists)? In many countries, antidiscrimination laws prohibit employment discrimination based on protected classes (e.g., on the grounds of race, skin colour, ethnic origin or gender). And it is arguably the case that just as selectors should not select candidates on the basis of such features, reaction qualifications grounded in customers' potential racist and/or sexist reactions are illegitimate.[25] In circumstances in which it is not allowed for the employer to reject the candidate based on costumers' racist or sexist prejudices, the disadvantages associated with not counting the relevant reaction qualifications seem to be a matter of bad brute luck. Accordingly, on the unfair disadvantages' view, it seems unfair that the employer should be disadvantaged by hiring the candidate (i.e. by not counting reaction qualifications based on features pertaining to such group memberships)—and perhaps compensating the employer is a reasonable response in such cases.

7 | CONCLUSION

We have shown that there is a discrepancy between, on the one hand, central components in a leading account of the (il)legitimacy of reaction qualifications and, on the other, our findings of people's moral evaluations of the legitimacy of key cases of reaction qualifications. Only one of the two components of the unfair disadvantages view generally mirror our respondents' intuitions. According to our findings, when the candidate's appearance feature does not reflect choices on their part, people are significantly less inclined to consider it okay to count the reaction qualifications against hiring the candidate. However, people's views on the appropriateness of counting the reaction qualifications against hiring the candidate are not affected (on average) by whether the relevant appearance feature reflects values or convictions of importance to the candidate. That is, apparently people's intuitions do not show sensitivity to a factor whose importance has been highlighted by the moral theory of reaction qualifications we have been considering.

Against the backdrop of these findings, we addressed the difficult question of how such information might impact our philosophical reflections. We point to two avenues in that regard: First, such findings may not warrant regarding a reaction qualification that is illegitimate from the philosophical perspective (one that is arguably superiorly reflective) as legitimate. However, it may appropriately motivate regarding it as *less* illegitimate. Second, without affecting the question of illegitimacy, the findings in question may affect the all-things-considered acceptability of letting the reaction qualification in question count against hiring the candidate. Specifically, our findings may contribute to making it overall acceptable to count the qualification (although the reaction qualification as such remains illegitimate).

In these ways, our article contributes new empirical insights to the political philosophical field of discrimination and reaction qualifications. Additionally, we make suggestions that appear to be of broader relevance, especially with regard to how laypeople's intuitions may appropriately figure in moral theory assessments.

ACKNOWLEDGEMENTS

We have incurred many debts in the form of helpful comments and suggestions in the process of writing this paper, both from fellows of the Centre for the Experimental-Philosophical Study

[25]Cf. Lippert-Rasmussen's symmetrical view (2014, pp. 245–246).

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/theo.12559 by VIVE – The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of Discrimination (CEPDISC) at Department of Political Science, Aarhus University and from others. We would like to thank especially Bjørn Gunnar Hallsson, Nils Holtug, Lasse Laustsen, Kasper Lippert-Rasmussen, Helene Helboe Pedersen and Michael Bang Petersen. We are also grateful to the participants at the CEPDISC 2021 conference for their useful feedback.

FUNDING INFORMATION

This work was supported by the Danish National Research Foundation under Grant DNRF144 and Society of Applied Philosophy under Grant SAPPDS112021DBA.

CONFLICT OF INTEREST STATEMENT

The authors report there are no competing interests to declare.

DATA AVAILABILITY STATEMENT

All data and scripts necessary to reproduce the results presented in the paper are available at Harvard Dataverse: https://doi.org/10.7910/DVN/WFKCPA.

ETHICS STATEMENT

The research material consists of survey data, which has been collected in direct engagement with the participants. Participation in the surveys was based on voluntary consent. The pool of participants was based on an approximate random sample and was therefore diverse. We considered the risk of harming any of the survey participants—including participants from groups that could be considered as vulnerable—as very low given the nature of the questions in the surveys (we did not consider the questions to be sensitive). For these reasons, and because ethical review of studies based on standard surveys are not required in Denmark, the surveys have not been reviewed by an ethical review board.

ORCID

Viki Møller Lyngby Pedersen https://orcid.org/0000-0001-9761-6546
Didde Boisen Andersen https://orcid.org/0000-0002-4039-3502
Søren Flinch Midtgaard https://orcid.org/0000-0002-8443-5137
Kim Mannemar Sønderskov https://orcid.org/0000-0002-3550-0772

REFERENCES

Albertsen, A., Hallsson, B.G., Lippert-Rasmussen, K. & Pedersen, V.M.L. (2023) Does harm or disrespect make discrimination wrong? An experimental approach. *Philosophical Psychology*, (online first), 1–26.
Andersen, D.B., Midtgaard, S.F. & Pedersen, V.M.L. (2023) Er diskrimination på baggrund af reaktionskvalifikationer tilladeligt? Et politisk teoretisk perspektiv informeret af borgernes vurdering. *Politica*, 55(4), 333–348.
Bloom, H.S. (1995) Minimum detectable effects: A simple way to report the statistical power of experimental designs. *Evaluation Review*, 19(5), 547–556.
Bóo, F.L., Rossi, M.A. & Urzúa, S.S. (2013) The labor market return to an attractive face: Evidence from a field experiment. *Economics Letters*, 118(1), 170–172.
Bunel, M. & Tovar, E. (2021) Profit, morality and discrimination. *Applied Economics*, 53(40), 4692–4712.
Cíbik, M. (2020) On the immorality of tattoos. *The Journal of Ethics*, 24, 193–206.
Cohen, G.A. (1988) *History, labour, and freedom: themes from Marx*. Oxford: Clarendon Press.
Cohen, G.A. (1989) On the currency of egalitarian justice. *Ethics*, 99(4), 906–944.
Cohen, G.A. (1998) Once more into the breach of self-ownership: Reply to Narveson and Brenkert. *The Journal of Ethics*, 2, 57–96.
Cohen, G.A. (2004) Expensive Taste Rides Again. In: Burley, J. (Ed.) *Dworkin and his critics: with replies from Dworkin*. Oxford: Blackwell Publishing, pp. 3–29.
Dworkin, R. (2000) *Sovereign virtue*. Cambridge, MA: Harvard University Press.
Feinberg, J. (1986) *Harm to self*. Princeton and Oxford: Oxford University Press.
Greene, J.D. (2015) *Moral tribes: emotion, reason, and the gap between us and them*. London: Atlantic Books.
Greene, J.D., Sommerville, R.B., Nystrom, L.N., Darley, J.M. & Cohen, J.D. (2001) An fMRI investigation of emotional engagement in moral judgment. *Science*, 293(5537), 2105–2108.

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/theo.12559 by VIVE – The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Hansen, R.S. & Midtgaard, S.F. (2011) Sinking Cohen's flagship—or why people with expensive tastes should not be compensated. *Journal of Applied Political Philosophy*, 28(4), 341–354.
Johnson, S.K., Podratz, K.E., Dipboye, R.L. & Gibbons, E. (2010) Physical attractiveness biases in ratings of employment suitability: Tracking down the 'beauty is beastly' effect. *The Journal of Social Psychology*, 150(3), 301–318.
Kagan, S. (2001) Thinking about cases. *Social Philosophy and Policy*, 18(2), 44–63.
Kahane, G. (2013) The armchair and the trolley: an argument for experimental ethics. *Philosophical Studies*, 162, 421–445.
Kauppinen, A. (2005) The rise and fall of experimental philosophy. In: Knobe, J. & Nichols, S. (Eds.) *Experimental philosophy*, Vol. II. Oxford: Oxford Scholarship Online, pp. 3–30.
Khaitan, T. (2015) *A theory of discrimination law*. Oxford: Oxford University Press.
Lindauer, M. (2020) Experimental philosophy and the fruitfulness of normative concepts. *Philosophical Studies*, 177, 2129–2152.
Lippert-Rasmussen, K. (2014) *Born free and equal? A philosophical inquiry into the nature of discrimination*. New York: Oxford University Press.
Mason, A. (2017) Appearance, discrimination, and reaction qualifications. *Journal of Political Philosophy*, 25(1), 48–71.
Mason, A. (2023) *Personal appearance, discrimination, and disadvantage: what's wrong with lookism?*. Oxford: Oxford University Press.
Midtgaard, S.F. (2024) X-phi and theory acceptance in political philosophy. *Res Publica* (forthcoming). https://doi.org/10.1007/s11158-024-09678-8
Miller, D. (1992) Distributive justice: What the people think. *Ethics*, 102(3), 555–593.
Monk, E.P., Jr., Esposito, M.H. & Lee, H. (2021) Beholding inequality: Race, gender, and returns to physical attractiveness in the United States. *American Journal of Sociology*, 127(1), 194–241.
Nuffield Council on Bioethics. (2017) *Cosmetic procedures: Ethical issues*. Report. https://www.nuffieldbioethics.org/publications/cosmetic-procedures
Olsaretti, S. (2004) *Liberty, desert and the market: a philosophical study*. Cambridge: Cambridge University Press.
Olsaretti, S. (2008) Debate: the concept of voluntariness—a reply. *The Journal of Political Philosophy*, 16(1), 112–121.
Rawls, J. (1995) Reply to Habermas. *The Journal of Philosophy*, XCII(3), 132–180.
Rawls, J. (1999) *A theory of justice*, Revised edition. Oxford: Oxford University Press.
Raz, J. (1986) *The morality of freedom*. Oxford: Clarendon Press.
Rhode, D.L. (2010) *The beauty bias: the injustice of appearance in life and law*. New York: Oxford University Press.
Scanlon, T.M. (1998) *What we owe to each other*. Cambridge, MA: Harvard University Press.
Segall, S. (2010) *Health, luck, and justice*. Princeton and Oxford: Princeton University Press.
Singer, P. (2005) Ethics and intuitions. *The Journal of Ethics*, 9, 331–352.
Stemplowska, Z. (2009) Making justice sensitive to responsibility. *Political Studies*, 57, 237–259.
Stemplowska, Z. (2011) Responsibility and respect: Reconciling two egalitarian visions. In: Knight, C. & Stemplowska, Z. (Eds.) *Responsibility and distributive justice*. Oxford: Oxford University Press, pp. 115–135.
Tersman, F. (2008) The reliability of moral intuitions: A challenge from neuroscience. *Australasian Journal of Philosophy*, 86(3), 389–405.
Wellman, C.H. (2008) Immigration and freedom of association. *Ethics*, 119(1), 109–141.
Wertheimer, A. (1983) Jobs, qualifications, and preferences. *Ethics*, 94(1), 99–112.
Wertheimer, A. (2006) Reflections on discrimination. *San Diego Law Review*, 43(4), 945–980.
Widdows, H. (2018) *Perfect me: beauty as an ethical ideal*. Princeton University Press.
Yuracko, K.A. (2006) Sameness, subordination, and perfectionism: toward a more complete theory of employment discrimination law. *San Diego Law Review*, 43(4), 857–898.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Pedersen, V.M.L., Andersen, D.B., Midtgaard, S.F. & Sønderskov, K.M. (2024) Reaction qualifications in the eyes of the people: An experimental-philosophical study based on US survey data. *Theoria*, 90(6), 624–642. Available from: https://doi.org/10.1111/theo.12559

17552567, 2024, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/theo.12559 by VIVE – The Danish Center for Social Science Research, Wiley Online Library on [01/12/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License