Queering social categorization: The effect of binary measurements on categorical perception and perception of binary gender ( provisional title)

**Overview & theoretical background**

What can proponents of gender equality and inclusiveness expect of language? Implicitly, quite a lot. This is why we have fights for gender reform. And we have seen that gender-fair language opens up the mind. Conversely, is it the case that binary language closes the mind? Binary language is dominant in both psychology and in the world. So what happens if we move past that?

Sex and gender are complicated multifaceted concepts which cannot be summarized as simple binaries consisting of women and men (Hyde et al. 2018). In psychological experiments, however, that is precisely how sex/gender are frequently operationalized. Recently there has been a growing awareness that how researchers ask participants about their own gender identity, solely providing the responses woman and man is limiting and alienating to many (Linqvist et al., 20018; Westbrook & Saperstein, 2015). There has yet to be a similar examination of how participants are asked to categorize others. Often these are based in normative views on gender, and may consequently bias participants in favor of those views. Here, I propose two studies to examine whether binary views on gender communicated through response options shape participants outcome. Study 1 examines whether binary response options enhance categorical perception of gender. Study 2 examines whether the explicit inclusion of queer response options increases queer gender categorization.

An important starting point for this study is queer theory, which is perhaps best understood as any method that begins by questioning the normative idea that gender, sex and sexuality are best understood as discrete and easily knowable categories (Ellis, Riggs & Peel, 2019). A key insights of queer theory is that sex/gender/sexuality consist of a network of ideas structured around a contrast between *normal, which* consists of heterosexual women and men and *Other(or deviant?)* which is anyone that does not neatly fit into these categories (Thorne & Hegarty, 2019). Consequently, the addition of non-normative categories still maintains the normal/Other distinction and queering a certain category has become shorthand for exploring the ways that category is mutable, flexible, and in which categories are not mutually exclusive. (Carr, 2015). Queer theory is a useful tool for challenging normative gender ideals, and can therefore function as as a lens to examine existing research and as a guide for future research (for example, Morgenroth et al., 2018).

Another key idea from queer theory is that gender is naturalized through repeated actions and through language. If gender does not result from physical bodies, where does it come from. Well, it’s performative. It comes into being through repeated acts that ascribe meaning onto bodies. Gender is performed and the act of the performance creates the concept. Peformativity and the role of language, concepts and norms. These are all constraints on ideas of sex/gender.

Recently, a number of studies have been published suggesting support for these ideas from queer theory. For example, Hyde (2019) examined a wide range of studies reporting gender differences between women and men and concluded that in many cases, the gender differences were exaggerated due to focus on average differences over within-category variation. Joel and colleagues (2014) showed people don’t always feel like their gender, even cis-people. Hester and colleagues examined ratings and features of faces and found no evidence that there are two distinct types of gender in faces. These studies suggest that normative assumptions of gender don’t always hold up. These studies show that queer theory can be a useful guid to research.

**Social categorization**

Social categorization research examines the process and consequences of mentally sorting others according to membership in various social groups (see Freeman & Anbady, 2011; Fiske & Taylor, 2013). Activating certain social categories, such as women or men, also activates associated ideas and stereotypes (Macrae & Bodenhausen, 2000; Freeman & Anbady, 2011). The research differentiates between thinking about someone categorically, where they are primarily appraised according to the social group they belong to or individualistically, where they are appraised as individual (Fiske et al., 1998). In the last few decades, researchers hav also started investigating social categorization as it is evoked by faces.

Although there is diversity in the type of effects they examine, many studies use what I will term the “categorization paradigm”, a method where researchers present participants with a face and ask them to categories it according to one of specified social categories. For gender, these categories are almost exclusively “woman” and “man” (or “male” and “female”) (see for example, Cloutier et al., 2005; Campanella et al., 2001; Webster et al., 2004; Zhao & Bentin, 2008). This paradigm has been use to show that participants categorize gender in faces even when gender was task-irrelevant and when not instructed to do so (Mason, Cloutier, Macrae) and when presented with images degraded to the point that no individual qualities could be discerened (Cloutier, Mason & Macrae, 2005). Together, these studies present a view of binary gender as a powerful organizing principle for humans.

**Toward a queer social categorization**

Although the consequences and content of social categorization are well studied, relatively less time is spent considering how people represent categories. Thorne and Hegarty (2019) draw from both cognitive psychology and queer theory to outline several possible ways represent categories. According to various systems of categorization, categories can be more or less discrete, more or less permeable and more or less mutually exclusive. The point is not that gender categories are necessarily one way or another –though some research suggests that they are probably fuzzy, permeable and mutually inclusive (see Hyde et al., 2018) – the point is that different participants may have different views on the nature of gender categories. In contrast, the categorization paradigm, where participants have to choose from a selection of possible options, presents a view of gender categories as discrete, impermeable, and mutually exclusive. Bringing it back to Butler, this is a performtive act. It recreates gender in a laboratory setting. Are participants influenced by these implicit messages? We don’t know! And on the other hand, does language that decenters binary gender also reduce the power of it as an organizing principle? Possibly!

Of course, this raises the question, what *does* a queer social categorization look like? It would not be categorizing someone *as queer*, rather, it would be allowing them to categorize into any number of not possibly co-existing identities. Measuring this type of categorization poses a challenge, however. One possibility is to supplement the traditional binary response categories with additional alternatives and the option to refrain from categorizing altogether. Indeed, when participants were given these options, many opted into them, suggesting that the standard categorization paradigm is unnecessarily restrictive (van Berlekom et al., in preparation). Such expanded categorization paradigm does not present gender as binary, but, it still implies that gender categories are mutually exclusive and discrete. The inclusion of queer categories additionally comes with the somewhat uncomfortable implication that queerness looks or should look a certain way. Despite this issues, an expanded social categorization paradigm is still more open than a simple binary categorization paradigm.

Another possibility for measuring queer categorization may be gender-neutral pronouns, such as the English singular “they”. “They” can be used to refer either to a person with a non-binary gender identity or to a person of unknown or irrelevant gender. This ambiguity means that to referring to someone using “they” leaves space for them to inhabit any possible gender identity, or multiple identities. In other words, it is queer. In contrast, using “she” or “he” is to implicitly mark someone out as belonging to the gender category associated with that pronoun. In other words, pronoun use is indicative of implicit categorization, and can therefore be useful as a measure of queer gender categorization. But it most be noted, is not necessarily a direct categorization in and of itself. Nevertheless, I think pronouns can still be useful for measuring the effect of how questions of gender affect queer social categorization.

**Categorical perception as a tool for measuring binary thinking**

Categorical perception could be a good measure of the extent to which binarity influences participants. Whaaat? Why? Well, I’ll tell you. Categorical perception is a visual bias where dimensional stimuli come to be seen as belonging to discrete categories. Such effects are well-observed for both vowel sounds and color, where small shifts in frequency or color spectrum result in clear shifts in color category or vowel sound. It is an important example of concept shaping perception.

The relevance here is that categorical perception has also been observed for faces. Using faces morphed on a continuum from feminine to masculine, multiple researchers have found that faces near the midpoint are catageorized as the endpoint categories more than would be expected just based on their features (Campanella et al., 2001; Freeman et al., 2012; Thorne e al., 2015). Importantly, categorical perception is flexible. Categorical percptions of individual identities can be learned (Ref) Categorical perception can be a function of individual beliefs (categorical perception of race). Categorical perceptions suggest that gender is viewed normatively.

Bring it to back to both T & H and Butler. The point of T& H is really that there a number of different ways to view categories. Based on phrasing, researchers doing gender research view it according to the *classical* view. This is a performance of gender, a recreation of gender. Does this classical viewpoint transmit to participants? We’ll use categorical perception to find out. The presence of categorical perception indicates people are seeing gender according to the classical view, the absence, that it does not.

**Overview of the present research**

The overarching goal of the study proposed here is to investigate whether categorization of gender measured with response options that challenge binary gender norms leads participants to view gender categorically and to categorize gender as binary or queer. A key difficulty in this research is that the outcome measure is dependent on the response options. In other words, it’s very easy to conflate the independent variable with the dependent variable. A categorization task with woman/man/I don’t know is simply not comparable to one with woman/man as options. I see two ways to overcome this problem. The first possible strategy is to find some other measure which is a consequence of the main outcome. This is the strategy used in study 1, where the outcome is not binary categorisation *per se* but categorical perception, as a kind of proxy. The second possible strategy is to fenagle two measures that are both different in terms binary options, but also directly comparable. This is the strategy used for study 2.

**Study 1**

Categorical perception will be tested using a categorization task with two conditions. Participants are shown a number of faces and asked to categorize them according to gender. In the *binary* condition, participant are given the options woman and man. In the *multiple categories* condition participants are given the options of categorizing according to the categories woman, man, non-binary and I don’t know. Participants in both conditions categorize 50/50 morphs and 60/40 morphs. The outcome measure, *Categorical perception* is calculated in both conditions as the average difference between the umber faces categorized as women for 50/50 and 60/40 faces in both conditions.

Hypothesis: There will be more categorical perception in the binary condition compared to multiple categories condition.

What would this result show? Categorical perception implies that participants view faces as being only women or men, and moreover, that women and men have two distinct sets of appearances.

**Study 2**

Social categorization will be tested using a slightly modified version of the classic categorization paradigm. Participants are presented with a face and then asked which pronoun they would use to refer to them. In the *multiple* *options* condition, participants choose from three possible pronouns: They, she or he. In the *free text* condition, participants are given a text box where they can write any pronoun the like.

Hypothesis: Participants categorizing according to multiple response options use “they” for more faces than participants responding with a free text.

Some stray thoughts:

**Snips**

I suggest a modification of Campanella et al’s categorization study with the important independent variable being be measurment type. Either binary woman/man or multiple options, including “I don’t know” and non-binary, although there could be room for others as well. Campanella et al. Found perception to be categorical, which meant that faces were classified as women more often than the percentage of dimorphic “femininity” would suggest. This is illustrated using the red lines in figure 1. If binary measurment is causing participants to exaggerate perceived gender differences, we might expect the pattern of result to be similar to the blue line in figure 1. In other words, a 70% man morph would be categorized as “woman” 100% of the time.

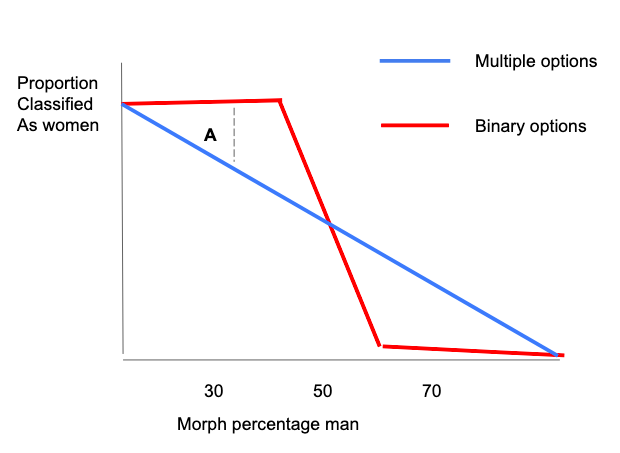


Figure 1. Hypothetical pattern of results from a categorization study comparing binary response options to multiple options. Dotted lines **A** epresent my suggestion for the main difference between the two conditions.

**Analyses & comparisons**

This presents a potentially difficult testing situation. There are a potential number of degrees of freedom. Which type of categorization should be compared? Which faces faces? I suggest a number of simplifications. First, I suggest limiting the analyses to one categorization, for example “women” categorizations.

Second, I see at least two ways to test the effect. One is illustrated by the dotted line **A** in Figure 1, simply comparing the number of 70/30 faces categorized as women when participants use binary or multiple response options. In this case, the relevant comparison would be of 70/30 faces only and other morph values may be included as within-subjects factors, but would only tested as negative controls because I would not expect there to an effect of response option at 50/50 faces or 100/0 faces. This would be quite simple. However, a possible concern is that the two conditions are so different that a direct comparison does not make sense.

Another option is illustrated in Figure 2 and would attempt to produce a measure which is more directly comparable. This option would require at least two morph types 70/30 and 50/50 as a within-subjects factor. The dotted lines B and C in figure 2 represent the average difference between 70/30 faces and 50/50 faces for participants responding using binary options (B) and multiple options (C). The relevant comparison would then be between the lines B and C. Comparing the difference scores would add a layer of complexity to the study. On the other hand, difference scores are arguably more comparable than the raw scores.

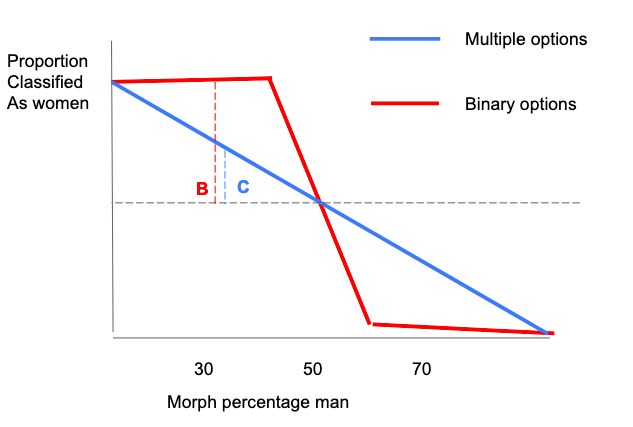


Figure 1. Hypothetical pattern of results from a categorization study comparing binary response options to multiple options. The Dotted lines **B** and **C r**epresent the hypothesized difference between 50/50 faces and 30/70 faces in each condition. The lines are meant to represent the same point on the X axis, but are kept slightly apart for illustrative purposes only.

**An alternative way to go**

I had another thought for a slightly different way to go as well. This idea is not clearly developed, but I’ll sketch it out very briefly. Basically, the thought is to investigate how measurement type (binary/multiple) affects which pronouns people use to describe a face. Here I’m imagining first you complete a categorization task similar to what we had in Study 1 and then people are either asked to describe the face or asked outright which pronoun they would use to describe that face. This gets at some conversations we’ve had earlier, pronouns are not labels, but they are indicative of categorization. In previous versions of the manuscript, we’ve talked about how misgendering is the real problem, not necessarily miscateorization, and what would be be best for most people would be if everyone avoided categorizing others, specifically verbally. Basically, someone can make an automatic categorization of someone else as “woman” but then have the second thought that gender is not visible from the face and use gender-neutral pronouns. I’m not sure if this is the best idea for the current study, but this sort of task could be a good continuation of study 1 in any case, with a measure that is closer to the relevant real-world behavior that we are interested in.

However, many of queer theories are compatible with psychology, and often the results show that the ways people categorize others are “queer” (Thorne & Hegarty, 2019).

Something about how experimenters bring their own assumptions about gender into studies. How

do psychologists talk about and treat gender categoreis? Well, very clearly as a binary, but also as a “natural” “essential” category. An important caveat to this research is that the binary nature of gender is almost invariably taken for granted in this research. Furthermore, gender is assumed by researchers to be an important organizing principle. For example,

Another challenge to this perspective is to point to the lived experiences of LGBTQ+ people. The existence of people who are Transgender and Gender Non-conforming (TGNC) suggests that the binary only studies are overly simplistic. Importantly, the queer experience is not defined by additional identities that function in a similar way to woman and man, but a multiplicity of flexible and fluid categories which may or may not be mutually exclusive to each other.

Much of the early work in this field evoked categories in a fairly simplistic way, by just writing out the category. This allowed that literature to examine ideas connected to social categories. Lately, we’ve started looking at faces too. The thing with faces is, they’re less digital, which more generally raises the question of what kind of categories are made. Furthermore, a face can be identified as belonging to a woman or a man, or it can be viewed as more or less feminine.

A big topic within this field is the degree to which gender categories are applied automatically. categorization. For gender specifically, there is some evidence from this body of work that categories are automatically activated. Researches in this tradition have identified gender as a particularly prominent characteristic and category, to the extent that people effortlessly judge gender from faces and automatically categorize others according to gender. Electroencephalographic (EEG) research has shown differential neural responses to feminine and masculine faces at 170 ms, leading to the common conclusion that people categorize gender within milliseconds of seeing a face (Freeman & Johnson, 2016; Kloth et al., 2010). Evidence for the automaticity of gender categorization comes from work showing that humans categorize gender even when not instructed to do so, (refs).

From this perspective, it also becomes quite clear that the literature of perception and categorization of gender makes a lot of assumptions based in the heterosexual matrix. Within this matrix is the idea that women and men are two distinct categories of people, with equally distinct appearances, and distinct facial features. In this system, faces are dimorphic and facial gender is defined as a single dimension with femininity on one end and masculinity at the other. According to this system, maleness, masculinity, and masculine facial features are all equally linked and femaleness, femininity, and feminine facial features are the opposite. The language used in studies supports this casual conflation of concepts.

I have a lot of underdeveloped ideas about measurment in general. I’m not even sure it will make it’s way into the final paper, but my thinking here is influenced by queer theory and feminist criticisms of science such as Haraway and Harding. One of the central ideas of queer theory is a rejection of the idea of categories. From a queer perspective, it becomes clear that the existence and importance of categories is routinely taken for granted by social psychology. Asking participants to categorize a face on the basis of two gender categories singled out by the researchers, from this perspective, constitutes not just a neutral measurement of a naturally occurring phenomenon, but a creation of a new situation where the researcher and subject work together to build a shared understanding of reality. The results from such studies are not necessarily invalid, but the queer perspective does show the literature has not considered the impact of measurement on the outcomes.

What does that mean for psychology? Well,this leads to the conclusion that how gender is measured can shape the results that we get. Therefore, we want to compare three types of measures of gender categorizations.

**Something something dimorphic perception of gender**

A related topic is the study of the dimorphic nature of faces and face perception. Gender as a facial characteristic is described as dimorphic, meaning as a single dimension defined as femininity on one end and masculinity on the other end. A related issue is whether are perceived dimoprhically. These are of course two different questions, but they go hand in hand Are faces dimorphic and are they perceived as dimorphic. Well, let’s see what the literature has to say.

The literature on face perception is a little bit complicated, but there are basically two approaches one can take when investigating the relationship between faces and perceived characteristics. The first is perhaps broadly defined as exploratory. Quantifying the shape of faces (using a field of research called morphometrics) and applying dimension reduction on that data. With this method, researchers have found one or two broad dimensions, which have alternatively been refereed to as approach and dominance (Todorov et al.,2008) and babyfacedness (Zebrowitz, 2007). The second approach is to investigate predefined dimensions, based on the average faces of existing groups of people, such as self-identified women and men. Researchers using this approach have reported that scores on this dimension are correlated with ratings of masculinity and femininity (Komori et al., 2011; Mitteroecker et al., 2015; Zaidi et al., 2019).

Here a paragraph outlining the results of Huart et al., 2015 and Wittlin et al., 2018. The point I want to make is similar to the paragraph in study 1 outline those papers, which is that people’s perceptions of gender can change depending on external stimuli. In other words, it is *queer* in the sense of not being stable.

Hester et al demonstrated very clearly that masculinity and femininity are not a singular dimension. Analysing data from ratings of faces of two databases, they showed convincingly that many of the assumptions of the binary viewpoint were not supported by empirical tests. These included the assumptions that both dimorphic facial features reflect and perceived femininity and masculinity reflect a single latent variable. This fits in with Hyde, 2019 about how research tends to focus on gender differences, obscuring similarities. Hester shows that people don’t necessarily view gender as dimorphic when asked in a way that let’s them express not that.

The conclusion I want to reach from the combination of the last two paragraphs it that perception of gender is not necessarily dimorphic, but it could become more dimorphic if the researchers communicate binary and dimorphic ideas of gender to the participants. Measuring perceived gender using a one-dimensional scale does suggest that femininity and masculinity are opposite. Therefore, measuring perceived gender as singe one-dimensional scale may make participants more prone to viewing gender as dimorphic.

Categorical thinking, as applied to humans, is a mode of thinking where individuals are primarily appraised according to the social group they belong to, rather than any individual qualities they may be possessed (Macrae & Bodenhausen, 2000). Although the idea has some support from neurological studies, in the main, categorical thinking has primarily been investigated as an outcome, and which ideas and categorizing evokes. For example, what characteristics are evoked by a masculine face or a masculine name. However, as we shall see, when looking at the categories evoked by faces, that is where we start having a problem.

x

This section needs one final paragraph to bring the point home. Bring it in to a more concrete level about response options and so on. Basically, something to connect to the method described below.

The literature has spent a lot of time thinking about what the consequences of a social categorization are and the what leads to social categorization, but somewhat less time thinking about the type of categories social categories. As Thorne and Hegarty point out, categories come in many shape and sizes. Categories themselves don’t have to be conflicting with queer, if they are allowed to be porous, flexible, and mutually exclusive. And here we see the limits of the standard categorization paradigm. The categorization paradigm is based on the normative gender view that gender categories are distinct, mutually exclusive and binary entities which are visible from faces. In these experiments, participants can’t possibly make queer categorization, they aren’t *allowed* to.