Objectification and Vision: how images shape our early visual processes[1]

Note: this is a pre-publication draft, please cite final version

Alice Roberts

Graduate Center, CUNY

Introduction

Objectification involves treating someone (a subject) as a thing (an object). For instance, treating a woman as a sexual object. 'Treating' someone in a particular way can involve actions, speech, writing, images or ways of seeing. In this paper I will discuss the relationship between the latter two, how images can lead to a way of seeing that treats women as sexual objects. My focus will be on a particular way of seeing that involves a viewer's early visual processes.[2] I will demonstrate that objectification can occur in early vision, right where perceptual processes meet the world.

The specific type of image discussed in this paper, which is prevalent in the mainstream media, presents a woman as having a blank or sexualised expression and

---

[2] I will define 'early visual processes' as processes that occur in V1 (the primary visual cortex), early in the feed-forward process (which involves projection from lower-level neurons in V1 to higher level neurons). This includes the selection of visual content and other visual processing that is largely automatic and modular. Edge detection is a paradigmatic example of an early visual process.

emphasises her body at the expense of her subjectivity.[3] Following Kathleen Stock, I will call these 'mind-insensitive images' (Stock, p.296). I will argue that mind-insensitive images can have an effect on a viewer's early visual processes, specifically how their visual system is predisposed to select visual content. This is a new contribution to accounts of sexual objectification, as it focuses on the *uniquely visual* effects of these mind-insensitive images. Uniquely visual effects are significant, as the functioning of the early visual system is largely automatic and subconscious – meaning that mind-insensitive images could have an insidious effect on the parts of a viewer's visual system not under conscious control.

Most accounts of sexual objectification either sideline images or present them as speech, missing these uniquely visual features.[4] For instance, Catharine MacKinnon (1983) and Rae Langton (2009) describe pornography as a speech-act.[5] Since pornography often takes the form of an image, Langton and MacKinnon consider these pornographic *images* as *speech*.[6] Examining visual pornography in a linguistic framework has led to some productive insights, such as suggesting that pornography can transmit messages, set standards and rank groups of people. However, as previously mentioned, speech is one way of 'treating', while images are another. Images are not linguistic representations; they are visual representations. Although insightful in some

---

[3] In this paper I will focus exclusively on images that treat women as sexual objects. However, my account generalises to images that treat other groups in a way that denies their subjectivity, such as racial minorities and people of other genders.

[4] Nussbaum (1995) does not spend much time on images, Papadaki (2010) does not mention images, similarly for Haslanger (2012), Eaton (2017) and Langton (2009). Stock (2018) and Eaton (2012) are exceptions to this.

[5] Mackinnon and Dworkin (1983), and Langton (2009) who follows their definition, define pornography as sexually explicit material. Although the images I am speaking about have sexualised content, they are not sexually explicit. They would not therefore be considered pornography by this definition.

[6] Nussbaum (1995), Dworkin (1989) and Bartky (1990) also do not distinguish between visual and linguistic representations.

ways, using a linguistic lens neglects and obscures important effects that images can have on a viewer in virtue of being visual representations.

One such important effect is the influence that repeated exposure to certain images can have on a viewer's early visual processing system. This effect will be the focus of my paper. It is important to note that focusing on the visual features of an objectifying image does not invalidate the significant insights provided by speech-act theory. Rather than arguing against a speech-act based account, I am suggesting that the nature of images *as visual representations* should be taken more seriously in accounts of sexual objectification.

This paper is divided into three sections. In the first, I will narrow my focus to one aspect of objectification: the denial of subjectivity. I will also describe how my account can build on a previous account of objectification offered by Kathleen Stock. In the second section, I will give a brief outline of 'visual priors', the early visual selection mechanism which I claim is affected by mind-insensitive images. In the third, I will consider how mind-insensitive images shape visual priors and argue that their effect is important to consider for an account of sexual objectification. I will also explore the pernicious epistemological and ethical consequences that this new form of objectifying seeing entails, as well as its consequences for correcting objectification and moral responsibility for objectification.

§1: How Can Images Deny Subjectivity?

Objectification can occur in many different ways. Nussbaum (1995, p.257) lists seven. The way of objectifying relevant to my argument is the denial of subjectivity. I will take

'subject' to mean human being, and 'subjectivity' to mean the expression of feelings, emotions and other psychological or mind-involving traits.

In this section I will argue, following Stock (2018), that images can be 'mind-insensitive', and also outline how my account can build on Stock's. This section will set up the initial grounds for my central argument; that if mind-insensitive images are viewed often enough and presented in certain contexts, they can cause viewers to see in a way that denies the subjectivity of individuals that fall under the type depicted in the image.[7]

§1.1: How can an Image be Mind-Insensitive?

According to Nussbaum (1995, p.257), denying subjectivity involves 'the objectifier treating a person as someone whose experience and feelings (if any) need not be taken into account', an example of which could be a manager staring at a woman's breasts during a job interview. The manager is neglecting the interviewee's qualifications, emotions and intelligence, instead seeing her as only a physically desirable body. This denies or ignores her position as an interview candidate, and more broadly her position as a minded subject.

As illustrated by this example, it is relatively clear how certain acts can deny subjectivity. By contrast, the way that an image can lead to the denial of subjectivity is more complicated. An image is not an action, so unlike the manager described above, cannot deny subjectivity by itself. Rather, I am going to argue that a mind-insensitive image has an influence in a circuitous way, through its *perceptual effects* on a viewer.

---

[7] For the purpose of this paper, I will assume that an image can depict a type (i.e. 'women') not just tokens, I will also assume an experiential account of depiction, which explains depiction in terms of the kind of experience an image causes in a viewer (Gombrich, 1960; Wollheim, 1987; Newall, 2011).

These effects result in the viewer becoming more likely to see women in a way that denies their subjectivity.

In order to have these effects, an image must first depict a woman in a way that does not affirm her subjectivity or mind; it must be a 'mind-insensitive image' (Stock 2018, p.296). The two most important features of a mind-insensitive image are:

1. Limited facial expression: The woman is presented as having a blank, expressionless face, or her facial expression is very limited, with the sole purpose of being sexually provocative (Stock 2018).

2. Foregrounding of erogenous zones: The woman is posed in a sexualised way to make breasts, crotch, and/or buttocks the focus. This also often acts to shift focus away from their face or other more subject-associated features.[8]

An image that contains both 1 and 2 is a Calvin Klein advert from 2016:

Figure 1

---

[8] Eaton (2012) speaks about the foregrounding of erogenous zones as characteristic of the female nude in canonical art. These features are also often seen in contemporary advertisements.

The image presents the woman depicted as having a blank expression, with her erogenous zones foregrounded. Furthermore, the flesh-tones present in the image and the way that the woman's legs and arms are opened towards the camera emphasise its sexual nature, drawing attention to the bodily, sensual features of the woman depicted, and away from her mind. Through possessing these features, the image can be considered a mind-insensitive image.

At this point, a sceptic may intervene and say that an image cannot depict a person's 'mind' or mental states. Resemblance theories of depiction claim that a depiction resembles what it depicts, for instance through having the same occlusion shape or outline shape (Hyman, 2006; Hopkins, 1998). Since a mind or mental state is not something that has an occlusion or outline shape, the sceptic could claim that it cannot be depicted. However, I would contend that the mental states themselves do not need to be depicted, rather they can be suggested, or made salient, through a combination of facial expressions and body language. Indeed, in real life, this is all we have to go off – we do not 'see' mental states – yet we manage to determine mental states fairly well. So, in images, as in real life, desires, thoughts, emotions and intentions can be suggested to the viewer through the facial expressions and body language of the person depicted. As a result, in what follows, I will adopt an experiential account of depiction, which explains depiction in terms of the kind of experience an image causes in a viewer (Gombrich, 1960; Wollheim, 1987; Newall, 2011). According to this account, mind-insensitive images will cause a viewer to experience a blank or sexual expression in the woman depicted and not be able to deduce much more about the depicted woman's mind. By contrast, in non-mind-insensitive images, the viewer will experience the depicted woman as having a more complex mental state.

To give a more concrete example, an image in which the mental states of the woman depicted are salient is this photograph by Dorothea Lange:

Figure 2

The woman is smiling slightly, with a hand on her hip, and one of her legs on the porch. Her expression conveys something like pride (perhaps that is her house in the background? Or she is enjoying her photo being taken?) but with an undertone of playfulness, quiet excitement and awareness that she is 'posing'. Like the Calvin Klein advert, this is also a low angle shot of a young woman, however, it does not display features (1) or (2). In this photograph, the woman's mind is clearly on display – as an audience, we wonder what she is thinking about to make her smile and place her hand on her hip proudly; we wonder who she is.

§1.2 Stock's Mind-Insensitive Seeing-As

Stock (2018, p.304) claims that repeated exposure to mind-insensitive images (such as Figure 1) can cause a viewer to develop a way of seeing that is mind-insensitive. This involves forming a perceptual gestalt of woman-as-body. The perceptual gestalt then projects onto real women based on aspects pulled out of these images. It results in a way of seeing women that simultaneously involves (a) attention towards bodily and other physical characteristics and (b) attention away from minded features (Stock 2018, p.302). She calls this mind-insensitive seeing-as.

Although Stock is arguing for a specifically visual effect, she remains vague about the level of seeing at which mind-insensitive seeing-as occurs. The verb 'seeing' can refer to many different levels in the process of vision: early visual processing, later perceptual attention-allocation and even later post-perceptual evaluation and inference. It is likely that her mind-insensitive seeing-as happens in the second or third stages, since she refers to some fairly complex perceptions.[9] Admittedly, it is not necessary to specify exactly where in visual processing the objectifying gestalt takes hold for Stock's argument to succeed. All she needs to show is that mind-insensitive seeing-as happens at *some* level of visual experience. In §3, I am going to propose a promising candidate for a specific visual mechanism (biased visual priors) that constitutes this way of seeing and specify where in visual processing objectification can occur.

Considering where in the visual process objectifying seeing occurs is epistemically and ethically significant. If objectifying seeing happens very early in visual processes, it is likely to be automatic and more difficult for a viewer to resist. Similarly, it would throw into doubt the idea that objectification can be countered by looking at, and neutrally seeing, a person as they really are. This is a significant insight for the epistemology of perception, as it sets out a new way that bias can occur before

---

[9] Such as seeing women-as-animals and other metaphorical perceptions.

perceptual knowledge is constructed, and is also significant for ethics, for instance in evaluating the moral culpability of objectifiers.

In the rest of this paper, in line with Stock, I will argue that a mind-insensitive image can cause its viewer to see individuals that belong to the type depicted in a way that denies their subjectivity. I will suggest that this 'way of seeing' can occur at the level of early visual processing, due to the ability of images to shape an individual's visual priors. I will also suggest the possible harms and implications for objectification occurring in early vision, where visual processes meet the world.

§2: How do Visual Priors Shape Early Vision?

In this section, I will outline in more detail what is meant by visual priors and describe how they can lead to visual bias. I will also examine some empirical data in favour of visual priors and suggest that images (as oppose to objects) have more of an effect on visual priors than is acknowledged in the philosophical literature.

§2.1 What are Visual Priors?

'Visual priors' are the name given to a selection mechanism that determines our earliest visual content. I am going to argue that mind-insensitive images can make these visual priors become biased to selecting in an objectifying way. I will call this purported selection process 'objectifying-input-selection'. However, in order to argue for this, the way that visual priors 'select' content must be explained in more detail.

Visual priors are needed because the visual system is exposed to more information than the brain can process. One way to solve this problem is for the visual system to

maximise the extraction of information from a given input by drawing on previously encoded information (Davenport and Potter 2004; Woodman et al. 2013).

Constructivist models of perception conceive of early visual processes as working this way: extracting some distal stimuli from the series of possible distal stimuli based on probabilistic predictions (Helmholtz 1867; Gregory 1970, 1980; Hochberg 1981). Visual priors encode information about past visual experiences and make a prediction that the regularities found in that sample can be projected onto the future (Gregory, 1970). They then select a distal stimulus based on these regularities. The selected distal stimulus is what becomes visual content. The process of priors being shaped to select a certain distal stimulus is called visual statistical learning.

This means that individuals with different experiences can end up with different priors. So, based on the modifications to an individual's visual priors by visual statistical learning, they could at a later point look at the same thing, but have different visual experiences. This is because their visual priors have adapted to pick out different things as being the most likely distal stimulus, resulting in different visual content.

A benign example of this would be seeing, at a glance, a loaf of bread in a kitchen as opposed to a brown, rectangular brutalist sculpture. It is far more efficient for the visual system to assume at a glance that the brown, rectangular object is bread than to suspend judgement or predict something else, as in most cases it will in fact be bread.

The way that perceptual priors are shaped by visual statistical learning is considered to be evolutionarily advantageous. It involves the visual system working in (what cognitive scientists would call) a normal way, using energy-efficient methods for extracting information.

A consequence of this is that previous visual experience can determine what distal stimuli visual priors will choose, and thereby determine early visual content. However,

this seemingly benign feature of the visual system can have not-so-benign consequences, as we will see in the next section.

§2.2 How Can Visual Priors Cause Perceptual Bias?

Munton connects the functioning of visual priors to perceptual bias. In a society that has racial and gender injustices, viewers are likely to be exposed to visual statistical regularities that occur as a result of these injustices. Munton suggests that this can result in a viewer's visual priors selecting a distal stimulus in line with these unjust statistical regularities. She calls priors that select in this way 'gerrymandered' priors. These are visual priors which depend on a manipulation of the viewers environment to be accurate (Munton 2017, p.28).

An example of a person with gerrymandered priors could be a child who has only ever been taken by her parents to manicured parks, where there are a lot of squirrels, but no rats. So, when the child sees a furry, scuttling creature cross her path, her priors are gerrymandered to select for 'squirrel' over 'rat'. However, the accuracy of the prior is contingent on a continuing manipulation (i.e. the child's exposure to squirrels). It is a localised regularity. This means it is liable to revert; most furry, scuttling animals are not squirrels.

In this sense, the child's visual prior is epistemically unsafe (Sosa 1999, p.142). It only codes for a small-scale statistical regularity, which requires the parent's intervention (i.e. taking the child to manicured parks, not alleyways) to continue. The child's current visual priors would give rise to false perceptions in a wider environment.

Unjust statistical regularities in a society can be seen as functioning like the child's exposure to squirrels. They lead to individuals' visual priors being 'gerrymandered'

towards the injustice. Since the gerrymandered priors select for past statistical regularities, this means that a person is less able to perceive novel information (like the child being less able to perceive rats), making them less perceptually skilled. This inability to perceive novel information would perpetuate, in a specifically visual way, the injustices of the society

§2.3: Empirical Data in favour of Visual Priors

Results consistent with the biased selection of visual priors have been demonstrated by Correll et. al (2007). When instructed to shoot a human target only when the target is armed, participants were much more likely to shoot when confronted with the image of a black man than a white man (Correll et al., 2007, 2015). They were therefore much more likely to see an ambiguous object in the hand of a black man as a gun (i.e. the man being armed).

These results could be interpreted to show that the participants' visual priors were gerrymandered to select 'black man' as a likely combination with 'gun', after perceiving an unjust statistical regularity between black men and guns.

When speaking about what would constitute this visual regularity, Munton often focuses on the perception of objects. She uses an analogy with learning to identify seedlings:

> The novice gardener has to inspect each plant sprout carefully to distinguish weed
> from seedling where the expert gardener has no such difficulty […] These effects
> rely in part on the direction of attention, but also on changes to lower level
> perceptual processing the visual system is capable of changing in response to

> previous tasks and encounters. […] what information is extracted from a particular scene depends on the system's prior encounters (Munton forthcoming, p.5)

However, if the main influence on visual priors is seen to be 'the systems prior encounters', this complicates the aforementioned interpretation of Correll's study. The participants in the study were students, many of whom are unlikely to have ever seen a black man with a gun, and some of whom may have never seen a gun at all. Therefore, most of the students would have been unlikely to have encoded a statistical regularity between the two based on their perception of real-world *objects*.

The most obvious expansion would be to include *images*. In the media black people are frequently depicted as violent or as victims of poverty, both of which are closely linked to gun crime.[10] There are also frequent depictions of black men holding guns in posters, in films and on television. As images are likely to have made up the majority of the student's perceptual encounters with black men holding guns, they are likely to be the main source of the encoded statistical regularities that give rise to the perceptions demonstrated by Correll.

However, although images are seen to be a likely candidate to account for these perceptions, it remains to be shown whether images can in fact influence a viewer's perception of objects in this way. In the next section I will argue in favour of this claim. I will also extend Munton's analysis of racial bias to objectification, to argue for the main thesis of this paper.

§3 Visual Priors and Mind-Insensitive Images

---

[10] Based on data collected from a sample of local Chicago TV news between 1993 and 1994, news stories related to crime involving black people were also four times more likely to include mug shots than news stories related to crime involving white people (Entman & Rojecki, 2000).

Munton's work on visual bias, alongside the work in perceptual bias more generally, has focused on racial bias (Siegel 2017; Gendler 2011). In this section I will attempt to extend some of these findings to apply to objectification. This will build on Munton's account in two ways: by examining objectification as a form of visual bias, and by examining the role that images as opposed to objects play in causing this bias. It will also build on Stock's account of mind-insensitive seeing-as by specifying a type of objectifying seeing that occurs at the level of early visual processing.

An important class of disproportionate media images are mind-insensitive images. In terms of the particular visual features of these images, I will focus specifically on one pairing: a female face with a blank or sexualised expression (Figure 1).[11] If visual priors can be demonstrated to encode for these mind-insensitive images, this could constitute a bias in early visual processes, and would mean that objectifying-input selection could occur.

There are some features of mind-insensitive images that make it initially plausible that they would be able to cause objectifying-input-selection. Firstly, all images of people are concrete, distorted representations. They are concrete because they present a single, stationary depiction of a subject. They are distorted because they are always selective in what they portray about a subject. This 'selection' is extreme in the case of mind-insensitive images, as they only show women as having blank or sexualised expressions, while women really have many facial expressions. Therefore, mind-insensitive images expose an individual's visual system to a biased view of women, setting up the possibility of objectifying-input-selection.

---

[11] There are other interesting features of images that I have left out, such as the context in which the image portrays the subject/object, its size and brightness, and the mood the image evokes, all of which will probably have some impact on what I am speaking about here. However, that is a topic for another paper.

Secondly, mind-insensitive images are plausible candidates for having an impact on visual statistical learning. The average London commuter is likely to be exposed to 3500 advertisements in a day (Story 2007). A significant proportion of these feature human models, and there is a strong correlation between female models and sexualised content.[12] Visual statistical learning is determined by the cumulative effect of perceived regularities. If a significant regularity between the depiction of women and sexualised content is perceived often enough, then it is likely that this regularity will be visually 'learnt'.

The fact that mind-insensitive images are a selective, sexualised portrayal of women, and so prevalent, offers some initial plausibility to the claim that such images can cause objectifying-input-selection. Also, visual priors operate on the level of 'types', not 'tokens': exposure to the individual stimulus object (e.g. a depiction of an individual woman) affects how further perceived individuals of that type (i.e. women in general) are processed by perception. This gives further initial support to the claim that viewing advertising images can result in acquiring priors that are relevant to the type 'women'.

However, to make a stronger case for this claim, our perception of images must be shown to be able to shape our perception of real people. Also, facial expressions and gender must be shown to be processed at the level of early visual processing. I will argue for these two statements in the next two subsections.

§3.1 Can Priors from Images Affect our Perception of Objects?

---

[12] Plakoyiannaki et al. (2008) collected data on female role stereotypes in online advertising. They took a sample of online adverts and classified them into eight different categories, some of which were gender-biased and others neutral. The two most frequent categories were women concerned with physical attractiveness and women as sex objects.

Recently, the convention in cognitive science studies to use images of objects as a substitute for objects themselves has been criticised. This is because recent empirical evidence has demonstrated that we have a richer visual experience looking at real objects than we have looking at images, like having different depth cues and other small effects that differentiate a 2D object from a 3D object (Dosso and Kingstone 2018).

However, there are also some important similarities between how we see images of objects and how we see objects themselves. The surface properties of images of objects are processed in a similar way to the surface properties of objects. Surface properties are things like colour and shape which are likely to be important in generating priors for the gender or facial expression of an individual. There is empirical evidence that the ventral stream (responsible for categorising objects through surface properties) is activated when both objects and images of objects are displayed (Grill-Spector et al. 2001). Processes in the ventral stream include processes in V1 and in other parts of the visual cortex. Since these parts of the visual cortex are relevant for acquiring priors, the fact that these areas are activated when viewers are exposed to both images and objects suggests that visual priors for surface properties can be acquired from viewing both images and objects.

Given these similarities in neural processing, our early visual processes are unlikely to distinguish between an object and an image when acquiring visual priors for surface properties. This is supported by the fact that, in practice, images serve important didactic purposes during early development. Children are often introduced to the visual appearances of unencountered objects through the images in picture books; they are taught to recognise objects through images. Overall, this suggests that while our early visual processing of depictions is not identical to the early visual processing of objects themselves, it is sufficiently similar in the relevant ways to acquire visual priors for

surface properties. This supports the claim that objectifying-input selection based on objectifying *images* of women, can be extended to *real-life* women.

§3.2: Can Priors Select for Gender and Facial Expressions?

For objectifying-input-selection to occur, gender and facial expressions must be shown to be content that is categorised at the level of early visual processing. However, making any claim about perceptual content is risky; it is a contested subject. While it is uncontroversial that simple, low-level properties like colour, shape, orientation and depth are processed at an early level, it is controversial whether other more complex properties are. Despite this controversy, recent evidence has given strong reason to believe that more complex properties called 'basic-level concepts', which includes the recognition of gender, are processed at this early level.

Basic level concepts are concepts like 'flower', 'dog', 'tree', 'gun', 'man' or 'woman'. These basic level concepts were confirmed experimentally by Rosch (1976). She demonstrated that when shown an image of an Alsatian, a person is far more likely to categorise the image as a 'dog' (basic level) than as an 'Alsatian' (more specific) or an 'animal' (less specific). Their availability for quick and natural categorisation is what makes these concepts basic-level.

There is empirical evidence for these basic-level concepts being processed at an early visual level. According to recent studies, visual categorisation of basic concepts can occur after an exposure of only 13ms (Potter et al. 2014; Keysers 2001). Only early visual processes function on such a short timeframe, so these studies provide evidence that the categorisation of basic-level concepts can happen during early visual

processing (Mandelbaum 2017, p.14). As previously mentioned, early visual processing is the level at which visual priors select information. Since the categorisation of basic-level concepts can be processed at this early visual level, and 'woman' is a basic-level concept, 'woman' is something for which visual priors select.

Determining the level at which facial expressions are processed is more difficult. However, the perception adaptation effects observed by Butler (2008) support the idea that facial expressions are processed at an early visual level. Perceptual adaptation effects occur in V1 (Blake et al. 2006), so if facial expressions can be shown to be subject to perceptual adaptation effects, it would entail that they are processed on this level too.

Butler (2008) demonstrates that when one stares at a face expressing anger for a minute or so, and then looks at an ambiguous face, the ambiguous face appears angry. There is evidence that this perceptual adaptation effect is not due to the low-level colour-shape properties of the face, but rather that the adaptation occurs relative to the more complex property of the facial expression. For instance, the effect persists across variation in low-level properties, as long as the facial expression is kept constant and faces that are inverted show a less significant adaptation effect than those that are upright (Butler et al. 2008). Arguably, this is evidence for a prima facie case that we have visual representations for facial expressions (Block, 2014), not just for the low-level properties that make up such expressions. Having visual representations for facial expressions suggests that the perceptual adaptation is based on directly perceiving the expression of 'anger' or 'fear'. Since adaptive effects have been shown to occur in early vision (Blake et al. 2006), and facial expressions (not just the low-level properties that comprise them) are subject to adaptation effects, this leaves an early visual level as the most plausible candidate for where facial expressions are processed.

Together, these findings suggest that facial expression and gender are processed in the early stages of visual categorisation. Therefore, both facial expressions and gender are types of content for which perceptual priors can encode statistical regularities.

As a result of this, if our environment demonstrates significant visual statistical regularity between female faces and blank or sexualised expressions, this could gerrymander visual priors towards selecting for visual content that shows female faces as blank or sexualised. This could then transfer from images of female faces to the faces of real women. The cases where this transfer is most likely is when a real woman has a facial expression that is ambiguous[13] – perhaps close to being blank (or sexualised), without actually being blank (or sexualised). In this sort of case, it is possible that a viewer with visual priors that strongly select for blank expressions will see the woman as having a blank expression, while a viewer who does not have these priors would not. In an individual instance, selecting for these expressions when confronted by a person who is not expressing them is epistemically and morally pernicious; it would involve attributing to that person mental states that they do not have.

There are differences between the two attributions I have focused on in this paper, namely, attributing a blank expression and a sexualised expression. Firstly, a blank expression displayed by an individual is usually interpreted by a viewer as the lack of a mental state in that individual. So, in usual circumstances, a viewer falsely attributing a blank expression to an individual wrongly suggests that the individual is lacking in mental states, and thereby denies that individual's subjectivity. By contrast, a sexualised expression, on its own, is not indicative of a lack of mental state. However, selecting disproportionately for sexualised expressions in women is an instance of the

---

[13] Another possibility that would make priors more likely to come into play is if viewing conditions are not ideal – e.g. if the woman's face is poorly lit, or partially occluded.

tendency to see 'women-as-bodies' described by Stock (2017, p.298). This tendency over-emphasises the bodies of women (in this case, their sexual desirability) and under-emphasises their minds, leading to an insufficient attribution of subjectivity. Therefore, an individual's visual system selecting for a blank or sexualised expression, when the individual is not expressing this, would be an objectifying way of seeing.

In terms of my overall project, the empirical material examined in this section gives support to the claim that perceptual priors are gerrymandered by their exposure to mind-insensitive images. Based on current empirical research, it is possible that priors in the early visual system select disproportionately for blank or sexualised facial expressions in real women, due to the way that facial expressions and basic categories are visually processed. As a result, objectification could occur at some of the earliest levels of visual processing, (i.e. the selection of visual content) right where the visual system meets the world. In the next section, I will discuss the epistemic and ethical consequences of this startling claim.

§3.3: Epistemic and Ethical Consequences

The upshot of considering the visual effects of mind-insensitive images is that such images (if seen often enough) could affect the selection processes of our early visual system. This is a significant contribution to the literature on objectification, as it gives some empirical support to the idea of objectifying seeing. It also specifies the level at which objectifying seeing could occur. In this way, objectifying-input selection builds on Stock's account of mind-insensitive seeing-as. It can also make clearer how the objectifying seeing caused by images can have bad epistemic and moral consequences.

However, before going into these consequences, it is important to pay more attention to the notion of causality I am using. I have postulated 'objectifying-input-selection' as a causal mechanism for the relationship between (A) and (B). However, note that mind-insensitive images (A) do not *by necessity* cause a way of seeing that is objectifying (B). Rather, they make an objectifying way of seeing more likely, through objectifying-input-selection, so (A) causes (B) in a defeasible, non-necessary way. However, even under this probabilistic sense of causation, objectifying-input selection can be morally and epistemically pernicious for the viewer, and objectifying image-making is a morally questionable practice to be engaged in.

Objectifying-input selection is epistemically pernicious because it is likely to lead to misperceptions. It involves selection of visual content based on objectifying images of women in the mainstream media, which do not track the expressions or phenomenological features of real women. As in the case of the child who sees all furry, scuttling animals as squirrels, the viewers of the media images have had their priors gerrymandered. The expressions and dominant presentation of women in mainstream media images are accurate only for these images; they encode a limited statistical regularity that is only true within the *domain of media images*. In other domains the statistical regularity does not hold. So, if objectifying-input-selection is transferred onto the faces of real women, as evidence suggests it can be, it can lead to false perceptions.

A tendency to see a group as lacking in subjectivity is also morally harmful, both for individuals in the objectified group and for the objectifiers themselves. I will examine these two perspectives in turn.

Firstly, seeing an individual in a particular group as having a blank expression means also a tendency to see them in a way that does not attend to their personhood.

People have a moral interest in having their personhood attended to, so not doing this can constitute a harm.

Women (considered as a group) having a good standing in society is something that has to be established, upheld, maintained and vindicated by society (Waldron 2012, p.60). A group's social status would clearly be threatened by a tendency for other individuals to deny their subjectivity and personhood. If my argument is correct, this can occur in a specifically visual way; if women's faces are indeed seen as sexualised or blank when they are in fact expressing different emotions, it would involve not upholding, maintaining or vindicating their good social standing, through not affirming their subjectivity. Furthermore, in addition to a tendency to misperceive women's facial expressions, women having a compromised social standing in society could have implications for their treatment by individuals in that society.

However, even if an objectifying way of seeing is not acted on, it could still be morally harmful to the objectifier themselves. Iris Murdoch gives the famous example of a mother seeing her daughter-in-law as flippant and in need of improvement. Internally, the mother-in-law is critical of every word and movement her daughter-in-law makes and sees her actions as expressing frivolity and loose morals. However, she is perfect in her outward demeanour towards the daughter-in-law, never saying anything dismissive or rude. Murdoch (2001) considers this way of seeing as stunting the mother-in-law's moral development, as it prevents her from seeing the positive aspects of her daughter-in-law. The mother-in-law can be faulted on both epistemic and moral grounds; she is missing out on knowledge about her daughter-in-law's attributes, and she is not respecting her as a human being. In order to grow morally the mother should reflect on her own prejudice and attend to or 'observe' her daughter-in-law differently, 'until gradually her vision […] alters' in a positive way (Murdoch 2001,

p.17). Moral improvement is considered by Murdoch to be a change in the mother-in-law's way of seeing or a cultivation of a form of attention that is virtuous. Therefore, having a biased way of seeing, or something that prevented her changing her way of seeing, would impede the mother-in-law's moral improvement.

Objectifying-input-selection could be a hindrance to a viewer changing their way of seeing women. In the case of objectifying-input-selection, a viewer's visual system is gerrymandered in a way that leans towards objectifying seeing, specifically, seeing women's faces as either blank or sexualised. This makes acquiring virtuous attention (which is required for moral development) likely to be much more difficult. Therefore, aside from being disadvantageous to the person who is seen in an objectifying way, objectifying-input-selection could stunt the moral development of the viewer by preventing them from cultivating virtuous attention towards women.

Paying attention to objectifying-input selection touches on the specific role images can play in a process of objectification. It demonstrates that they can lead to a way of seeing that is pernicious in both moral and epistemic terms, while also acknowledging the specifically visual features of these images.

My view also has implications for the possibility of correcting objectification and for an individual's responsibility for objectification.[14] The implication for correction is that the only feasible means of correcting objectifying-input-selection is a reduction in the number of mind-insensitive images on display. If I am right in postulating this link between mind-insensitive images and objectifying-input-selection, there is not much a

---

[14] These implications about correction and responsibility are only relevant to the form of objectification that I am discussing (i.e. objectifying-input selection). Objectification can occur in many ways and through many vectors, such via cognitive processing. These different forms of objectification will have different implications about correction and responsibility. For instance, objectification that is more on the cognitive side will (most likely) have greater scope for correction and attribute greater moral responsibility or blameworthiness to the objectifier. I take no position on whether the implications for correction and responsibility given by objectifying-input selection affect those same implications for other forms of objectification.

viewer can do to prevent visual-statistical learning from occurring or consciously reverse their objectifying seeing. If objectification occurs at the level of early visual processes, then just as viewer cannot 'see' the lines in the Muller-Lyer illusion as the same length, they also cannot 'see' in a way that is not objectifying, due to the modularity of early visual processes.

This leads on to the second implication: that a viewer is not morally blameworthy for the form of objectification I've been discussing. This is partially due to the fact that the content one's perceptual priors selects is not under conscious control. Similarly, viewers of objectifying media images, in most cases, do not produce these images themselves. Rather, it is corporations and advertising agencies that produce them. However, ascribing moral blame to corporations and advertising agencies for producing the images is perhaps simplistic, as they claim to produce images based on popular demand, which seemingly brings blame back to the viewers.

Due to the difficulty in ascribing moral blame, and the number of actors involved (corporations, photographers, advertising executives, viewers etc.), I will tentatively suggest that the best framework in which to examine responsibility for objectifying-input-selection is that of structural injustice, which abandons the notion of blame altogether. Structural injustices involve a dispersed set of individuals, who contribute to structural processes that end up giving rise to harms to other individuals (Young 2012). In this case, the 'structural processes' would be the production of mind-insensitive images, the contributors would be corporations, photographers, advertising executives, viewers etc., and the 'harm' would be the objectifying way of seeing women (i.e. objectifying input selection) caused by mind-insensitive images.

Despite not being blameworthy, individuals who are contributors to the processes that give rise to mind-insensitive images could have a political responsibility to rectify

the harms that occur. For instance, viewers of mind-insensitive images could have a political responsibility (Young 2012) to organise to stop corporations from producing or displaying the images, in order to prevent the pejorative influence these images have on how women are seen. This would not be a form of responsibility that ascribes blame to an individual viewer, but rather one that involves the responsibility to rectify a scenario that causes harm, in which the individual viewer is a participant (Young, 2012). However, these final statements are only meant as tentative suggestions - examining the relationship between objectification and structural injustice is a subject for another paper.

§4: Conclusion

The ability of mind-insensitive images to have an effect at the level of early visual processing demonstrates the significance of considering such images as visual representations, in addition to considering them as speech.

Previous literature on objectification has not considered empirical work on perception, despite the fact that objectification is almost always perceptually mediated. This paper has been an attempt to extend the findings of empirical work on racially biased perceptions to objectifying perceptions. The significance of this empirical work is to suggest that objectification can occur at the level of early visual processing.

Acknowledging the uniquely visual influences of mind-insensitive images and the significance of objectifying-input-selection brings new moral and epistemic consequences to the concept of objectification. This paper has explored these consequences, namely that objectifying-input selection can cause misperceptions, threaten the status of women in society and stunt moral development in viewers. Lastly,

the paper examined the implications that objectifying-input-selection has for correcting objectification and for responsibility for objectification.

My aim overall was to argue for a previously unconsidered influence that objectifying images can have when they are considered as visual representations. This influence is the ability of mind-insensitive images to cause women to be seen as lacking in subjectivity by affecting the early visual processing of viewers. This paper contributes a new angle from which objectification can be investigated, by highlighting the role of early visual processes; it shows how early visual biases can cause objectification right where our perceptual processes meet the world, and also demonstrates the epistemic and moral consequences of this new form of objectification.

Bibliography

Blake, Randolph et al. "Strength of early visual adaptation depends on visual awareness." *Proceedings of the National Academy of Sciences of the United States of America* vol. 103,12 (2006): 4783-8. doi:10.1073/pnas.0509634103

Block, Ned, "Seeing-As In Light of Vision Science". *Philosophy and Phenomenological Research*, 89 No.3. (November 2014): pp.560-572.

Butler, Oruc, Fox and Barton. "Factors contributing to the adaptation aftereffects of facial expression." *Brain Res*. 2008;1191 pp.116-126 doi:10.1016/j.brainres.2007.10.101

Cikara, Eberhart and Fiske, "From Agents to Objects: Sexist Attitudes and Neural Responses to Sexualized Targets". *Journal of Cognitive Neuroscience* 23(3). (March 2011): pp.540-551.

Correll, Joshua, Matthew Crawford, Bernd Wittenbrink, and Melody Sadler. "Stereotypic Vision: How Stereotypes Disambiguate Visual Stimuli." *Journal of Personality and Social Psychology* 108, no.2 (2015): pp.219–233.

Correll, J., B. Park, C.M. Judd, and B. Wittenbringk. "The Influence of Stereotypes on Decision to Shoot." *European Journal of Social Psychology* 37 (2007): pp.1102–1117.

Davenport, J.L. and M.C. Potter. "Scene consistency in object and background perception". *Psychological Science* 15 (2004): pp.559-564

de Jong, Maartje C., Jan W. Brascamp, Chantal Kemner, Raymond van Ee, Frans A. J. Verstraten. "Implicit Perceptual Memory Modulates Early Visual Processing of Ambiguous Images" *Journal of Neuroscience*, 34 no.30 (23 July 2014): pp.9970-9981 doi: 10.1523/jneurosci.2413-13.2014

Delgado, Richard and Stefancic, Jean. "When Hate Goes Tangible" in *Understanding Words that Wound*. New York: Routledge, 2004

Dosso, J.A. and A. Kingstone. "Social modulation of object-directed but not image-directed actions". *PLoS ONE* 13 no.10 (2018) https://doi.org/10.1371/journal.pone.0205830

Eaton, Anne. "What's wrong with the Female Nude?", in *Art and Pornography: Philosophical Essays*, edited by Hans Maes & Jerrold Levinson. Oxford: Oxford University Press, 2012. pp.287-291

Eberhardt, Jennifer L., Phillip Atiba Goff, Valerie J. Purdie, and Paul G. Davies. 2004. "Seeing Black: Race, Crime, and Visual Processing." *Journal of Personality and Social Psychology* 87 no.6 (2004): pp.876–93. https://doi.org/10.1037/0022-3514.87.6.876.

Entman, Robert M, and Andrew Rojecki. *The Black Image in the White Mind: Media and Race in America*. Chicago: University of Chicago Press, 2000

Flounders, Matthew W, Carlos González-García, Richard Hardstone, & J Biyu He. "Neural dynamics of visual ambiguity resolution by perceptual prior." *eLife*. 8. (2019) doi:10.7554/eLife.41861.

Gendler, Tamar Szabó. "On the Epistemic Costs of Implicit Bias." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, vol. 156, no. 1, 2011, pp.33–63.

Gombrich, Ernst H. *Art and Illusion*. Princeton: Princeton University Press, 1960.

Gregory, R. *The Intelligent Eye*. London: Weidenfeld and Nicolson, 1970.

Gregory, R. L. "Perceptions as Hypotheses". *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 290(1038), (1980) pp.181–197.

Grill-Spector, Kalanit, Zoe Kourtzi and Nancy Kanwisher. "The lateral occipital complex and its role in object recognition." *Vision Research: Elsevier*, (May 2001)

Haslanger, Sally. "On Being Objective and Being Objectified" in *Resisting Reality*. Oxford: Oxford University Press, 2012. pp.36-82

Helmholtz, H. Von (1867) *Treatise on Physiological Optic III*. The Optical Society of America, 1925.

Hochberg, Julian. "On cognition in perception: Perceptual coupling and unconscious inference". *Cognition* V.10, Issues 1–3, (July–August 1981): pp.127-134

Hopkins, Robert. *Picture, Image and Experience: A Philosophical Inquiry*, Cambridge: Cambridge University Press, 1998.

Hyman, John. *The Objective Eye: Color, Form, and Reality in the Theory of Art*. Chicago: University of Chicago Press, 2006.

Kulvicki, John. *Images*. London: Routledge, 2014.

Lamarre, Carl. "Drake Claims his Throne at New York MSG." *Billboard*. (8/25/2018) URL=https://www.billboard.com/articles/columns/hip-hop/8472171/drake-claims-throne-new-york-city-madison-square-garden-msg-migos-ovo-concert-tour Accessed=01/06/19

Langton, Rae. "Speech Acts and Unspeakable Acts." *Philosophy & Public Affairs* Vol. 22, No. 4 (Autumn, 1993): pp. 293-330

MacKinnon, Catherine. *Only Words*. Cambridge: Harvard University Press, 1993.

Mandelbaum, Eric. "Seeing and Conceptualizing: Modularity and the Shallow Contents of Perception." *Philosophy and Phenomenological Research* (2017) doi: 10.1111/phpr.12368

Munton, Jessie. "Perceptual Skill and Social Structure" *Philosophy and Phenomenological Research* (24 November, 2017) doi:10.1111/phpr.12478

Munton, Jessie. "Bias in a Biased System" In N. Ballentyne and D. Dunning (eds.) *Bias, Reason and Enquiry: New Perspectives from the Crossroads of Epistemology and Psychology.* Oxford University Press, (forthcoming).

Murdoch, Iris. *The Sovereignty of Good*. New York: Routledge, 2001

Nanay, Bence. *Aesthetics as Philosophy of Perception*. Oxford: Oxford University Press, 2016.

Newall, Micheal. *What is a Picture? Depiction, Realism, Abstraction*. New York: Palgrave Macmillan, 2011.

Nussbaum, Martha C. "Objectification". *Philosophy and Public Affairs*. Vol. 24, No. 4 (Autumn, 1995): pp.249-291

Papadaki, Lina. "What Is Objectification?" *Journal of Moral Philosophy* v.7 (2010): pp.16–36

Pierce, Charles-Saunders. *How to Reason: A Critick of Arguments*. MS [R] 397, 1893 URL= http://www.commens.org/bibliography/manuscript/peirce-charles-s-1893-how-reason-critick-arguments-advertisement-ms-r-397 Accessed=10/06/19

Plakoyiannaki, E., K. Mathioudaki, P. Dimitratos, et al. "Images of Women in Online Advertisements of Global Products: Does Sexism Exist?" Journal of *Business Ethics* 83 no.101. (2008): https://doi-org.ezp.lib.cam.ac.uk/10.1007/s10551-007-9651-6

Potter, M., Wyble, B., Hagmann, C., & McCourt, E. "Detecting meaning in RSVP at 13 ms per picture." *Attention, Perception, & Psychophysics* 76(2) (2014): pp.270–279.

Rosch, Eleanor, et al. "Basic Objects in Natural Categories." *Cognitive Psychology* 8 (1976): pp.382-439

Siegel, Susannah. *The Rationality of Perception*. Oxford: Oxford University Press, 2017.

Sosa, Ernest. "How to Defeat Opposition to Moore." Noûs, 33 s13 (1999): pp.141–153. doi:10.1111/0029-4624.33.s13.7

Sorrow, April Reese. "Magazine Trends Study Finds Increase in Advertisements Using Sex." *UGA U of Georgia*, (6 May 2015):

URL=http://news.uga.edu/releases/article/magazine-trends-study-finds-increase-in-advertisements-using-sex/

Stock, Kathleen. "Sexual Objectification, Objectifying Images, and Mind-Insensitive Seeing-As" in *Evaluative Perception*. Oxford: OUP, 2018.

Story, Louise, "Anywhere the Eye Can See, It's Likely to See an Ad" New York Times, (Jan 15, 2007) quoting statistics from Yankelovitch 'Market Receptivity' URL=https://www.nytimes.com/2007/01/15/business/media/15everywhere.html

Vadas, Melinda. "Could Pornography be the Subordination of Women?" *The Journal of Philosophy* Vol. 84, No. 9 (September 1987): pp. 487-511

Waldron, J. *The Harm in Hate Speech*. Cambridge: Harvard University Press 2012.

Wollheim, Richard. *Painting as an Art*. Princeton: Princeton University Press, 1987.

Woodman, G. F., Carlisle, N. B., & Reinhart, R. M. "Where do we store the memory representations that guide attention?" *Journal of Vision* 13 no.3 (2013)

Young, Iris Marion. Responsibility for Justice