European Journal of Analytic Philosophy EuJAP | 2025 | Vol. 21 | No. 1 | 39-55
https://doi.org/10.31820/ejap.21.1.3

SOME CRITICAL THOUGHTS ON "HOW STEREOTYPES DECEIVE US"

Federico Picinali[1]

[1] London School of Economics and Political Science, United Kingdom

Discussion – Received: 01/07/2023 Accepted: 01/11/2023

This paper is part of a book symposium on Katherine Puddifoot's *How Stereotypes Deceive Us* guest edited by Marina Trakas (IFF, CONICET)

ABSTRACT

The goal of this paper is to identify and discuss the weaker aspects of some of the arguments in Kathy Puddifoot's fascinating and thought-provoking book. Section 1 deals with Puddifoot's treatment of the "single factor view" and the "dual factor view" of stereotyping. Section 2 deals with Puddifoot's treatment of egalitarian attitudes. Section 3 deals with Puddifoot's treatment of the moral encroachment approach to stereotyping. Finally, section 4 deals with Puddifoot's theory of evaluative dispositionalism. The sections can be read independently.

Keywords: stereotyping; egalitarian attitudes; moral encroachment; pragmatic encroachment; evaluative dispositionalism.

Correspondence: f.picinali@lse.ac.uk

Introduction

The goal of this paper is to identify and discuss the weaker aspects of some of the arguments in Kathy Puddifoot's book, such that Puddifoot and other scholars interested in stereotypes may improve, at least marginally, on these arguments. Whether or not my criticisms are persuasive, it would be unfair to forgo the praise that Puddifoot deserves for her important work.

The book is replete with thought-provoking ideas, hypotheticals and arguments, making it a very valuable read for anyone interested in epistemology, and social epistemology in particular. However, there are two contributions that, in my view, stand out for their significance. The first, offered in Chapter 6, is the compelling analysis of the dilemma represented by an individual's choice whether to disclose their mental health condition and/or their social identity in certain contexts. Disclosure increases the risk of being stereotyped and of triggering a series of epistemic pitfalls of stereotyping, detailed by Puddifoot in earlier chapters, that will likely damage the individual facing the dilemma; non-disclosure may well mean that the interlocutor will misunderstand the needs of this individual, as well as their behaviour and attitudes. Puddifoot highlights a genuine problem that warrants further scrutiny. The brief discussion of strategies to tackle the problem (see, in particular, Puddifoot 2021, 128-132) is a valuable starting point for future inquiry.

The second contribution that stands out consists in the claim, defended in Chapters 7 and 8, that the rationality of holding a stereotyping belief *also* depends on the dispositions[1] that are possessed due to holding such a belief, or "downstream" dispositions. As pointed out below, I am not entirely convinced by the way in which Puddifoot defends this claim. The claim itself, though, is intriguing, plausible, and potentially disruptive of the mainstream views on the justification of beliefs.

I now turn to the criticisms, which are grouped in four distinct sections. Section 1 deals with Puddifoot's treatment of the "single factor view" and the "dual factor view" of stereotyping. Section 2 deals with Puddifoot's treatment of egalitarian attitudes. Section 3 deals with Puddifoot's treatment of the moral encroachment approach to stereotyping. Finally, section 4 deals with Puddifoot's theory of evaluative dispositionalism. The sections can be read independently.

[1] Dispositions are defined by Puddifoot as "what a person does, says, thinks, and would do and think in various circumstances" (2021, 164).

1. The single and the dual factor views of stereotyping

A first set of problems pertains to Puddifoot's critical discussion of what she calls the "single factor view" and the "dual factor view" of stereotyping. Both views concern the conditions under which stereotyping increases the chances that the agent makes an accurate judgement about a member of the group to which the stereotype refers.

According to the single factor view

> [T]here is only one feature of any act of stereotyping that determines whether the application of the stereotype (…) increases (…) the chance of an accurate judgement being made: whether or not the stereotype that is applied reflects some aspects of reality. (2021, 32)

On the same page, Puddifoot states that a stereotype

> [R]eflects some aspect of social reality as long as there is a regularity found within society and the stereotype leads a person to respond in a way that reflects the regularity. (Ibid.)

To show that this view is fallacious, Puddifoot offers the example of the stereotype associating Black people more strongly than non-Black people with drug use.[2] According to Puddifoot, this stereotype reflects an aspect of US social reality, this being the high arrest rate for suspected drug use[3] amongst Black people (ibid. 45, 61). And yet, Puddifoot says, the stereotype may also lead "to judgements that fail to fit accurate statistical information about actual rates of *drug use*, which are similar across Black and non-Black populations" (ibid., 45, emphasis in the original). This would allegedly show that the single factor view is wrong:

> An act of stereotyping might involve the application of a stereotype that reflects some aspect of social reality, which would mean that on the single factor view it should increase the chance of an accurate judgement being made. However, because the stereotype does not dispose the person who engages in the stereotyping to respond in a way that reflects the

[2] At pages 45 and 61 Puddifoot phrases the stereotype as referring to "White", rather than "non-Black" people. But in her discussion of the example, she switches between the two attributes. Moreover, the stereotype is sometimes phrased by Puddifoot as referring to "drug crime", sometimes as referring to "drug use". I have attempted to bring consistency to the example. Also consider the following footnote.

[3] NB: in England and Wales the law does not punish the "use" of drugs *per se*. Rather, the Misuse of Drugs Act 1971 targets conducts such as the importation, exportation, production, possession, and supply of controlled drugs.

> statistical reality, the stereotyping could reduce rather than enhance the chance of an accurate judgement being made (2021, 45).

To clarify, the stereotype in question may lead someone to conclude that a Black acquaintance is more likely than a non-Black acquaintance to use drugs. This conclusion—Puddifoot suggests—reflects the high arrest rate for suspected drug use amongst Black people; but it does not reflect the "statistical reality" concerning drug use.

My objection to Puddifoot's analysis is that the example she relies on does not show that the single factor view is wrong. Contrary to what is argued by Puddifoot, the stereotype about comparative drug use does not reflect the aspect of social reality represented by arrest rates, such that the single factor view would not insist that reliance on it increases the chances of making an accurate judgement concerning drug use. In other words, this view would likely lead to the same conclusion reached by Puddifoot about the accuracy of a judgement that relies on the stereotype.

To begin with, it is not clear in what sense the stereotype in question can be said to reflect the high arrest rate concerning Black people, given that the former is comparative (as all stereotypes are, according to Puddifoot's definition, to which I'll soon return), whereas the latter does not provide comparative information. The (spurious) proposition that Black people are more likely than non-Black people to use drugs is consistent with any arrest rate concerning Black people alone—insofar as this is higher than the rate amongst non-Black people—and may, therefore, "lead a person to respond in a way that reflects" a very low arrest rate amongst Black people. I take it, then, that with "high arrest rate" Puddifoot actually means that the arrest rate is *higher* for Black people than for non-Black people. Even so, the stereotype in question cannot, in fact, be said to reflect this comparative arrest rate.[4]

The stereotype may well lead someone to conclude that a Black acquaintance is more likely than a non-Black acquaintance to use drugs. Pace Puddifoot, though, this judgement about the comparative likelihood of drug use would not *reflect* the higher arrest rate for suspected drug use amongst Black people. The same holds for the stereotype on which the judgement is based. The judgement (and the stereotype) could meaningfully be said to reflect this comparative arrest rate only under the assumption that the comparative arrest rate tracks the comparative rate of

[4] By "comparative rate" I mean a construct that compares the rate for one group with the corresponding rate for the other group.

drug use, such that an inference could be reliably drawn between the two. To see why, consider the case in which the comparative arrest rate does not track the comparative rate of drug use, perhaps, because, due to racism in law enforcement, the arrest rate amongst Black people for suspected drug use is substantially higher than the rate of drug use amongst Black people. In such a case, no generalisation or individualised judgement about the comparative likelihood of drug use of Black and non-Black persons could reliably be formed based on the comparative arrest rate. After all, this rate is disconnected from the state of affairs about drug use, being determined, instead, by forces such as hatred, dislike, and suspicion towards Black people. If, however, the comparative arrest rate does not provide us with information on which to base reliably a generalisation or an individualised judgement about the comparative likelihood of drug use, it is mystifying to claim that the generalisation or judgement *reflect* the rate. They could reflect the comparative arrest rate only under the assumption that arrests are evidence of drug use. But this assumption is, *ex hypothesi*, false.

In the circumstances of Puddifoot's example, in fact, the assumption of a correspondence between the comparative arrest rate and the comparative rate of drug use cannot apply. In societies such as the US and England and Wales, where structural racism is a significant issue and where police forces are affected by institutional racism,[5] Black people are overrepresented among those arrested, both in general and with regard to drug-related offences considered separately. What is more important, a Black individual is significantly more likely than a White individual to be arrested, [6] notwithstanding that "BAME groups are less likely to commit crime".[7] Given the disconnect between the comparative arrest rate and the comparative rate of drug use, it is not clear how any judgement (or

[5] As far as the Metropolitan Police Service is concerned, see Baroness L. Casey, *Final Report. An Independent Review into the Standards of Behaviour and Internal Culture of the Metropolitan Police Service* (2023), available at https://www.met.police.uk/SysSiteAssets/media/downloads/met/about-us/baroness-casey-review/update-march-2023/baroness-casey-review-march-2023a.pdf (accessed 9 August 2023).

[6] For data about England and Wales see Ministry of Justice, *Black, Asian and Minority Ethnic Disproportionality in the Criminal Justice System in England and Wales* (2016), in particular, at 12, 22, available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/639261/bame-disproportionality-in-the-cjs.pdf (accessed 9 June 2023) and D. Lammy MP, *The Lammy Review: Final Report. An Independent Review Into the Treatment of, and Outcomes for, Black, Asian and Minority Ethnic Individuals in the Criminal Justice System* (2017), available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/643001/lammy-review-final-report.pdf (accessed 9 June 2023). For data about the US see Bureau of Justice Statistics, 'Race and Ethnicity of Violent Crime Offenders and Arrestees, 2018' available at https://bjs.ojp.gov/content/pub/pdf/revcoa18.pdf (accessed 9 June 2023) and R. Camplain and others, 'Racial/Ethnic Differences in Drug- and Alcohol-Related Arrest Outcomes in a Southwest County From 2009 to 2018' (2020) 110 American Journal of Public Health 85.

[7] Ministry of Justice, n 6, at 12.

stereotype) about the comparative likelihood of drug use could be said to *reflect* the comparative arrest rate; and it is not clear why a defender of the single factor view would maintain that said judgement (or stereotype) reflects said rate.

If the single factor view is indeed wrong, then, the example of a stereotype concerning the comparative rate of drug use amongst the Black and the non-Black populations does not contribute to show this. This is not an example in which the single factor view would insist that the stereotype reflects a relevant aspect of social reality (other than racism, that is). Notice, instead, that if the stereotype actually reflected the comparative arrest rate—due to the fact that the comparative arrest rate actually tracks the comparative rate of drug use—it would be perfectly sensible to argue that reliance on such a stereotype increases the chances of judging accurately whether a Black acquaintance is more or less likely to use drugs than a non-Black acquaintance. In this case the stereotype would indeed reflect the "statistical reality" concerning drug use.

Consider now the definition of the "accuracy" of a stereotype, introduced by Puddifoot in her critical discussion of another account of stereotyping, the dual factor view. Puddifoot states that

> A stereotype can be deemed to be accurate if it leads a person to respond in a way that is fitting with accurate statistical information about the distribution of traits across groups. (2021, 40)[8]

According to this definition, the accuracy of a stereotype does not reside in the accuracy of the generalisation that the stereotype represents: it consists, instead, in the accuracy of the judgement that is produced relying on the stereotype. This definition, though, seems to turn the dual factor view into tautology, since this view is construed by Puddifoot as claiming that

> [J]udgements produced as a result of acts of stereotyping are more likely to be accurate than alternative judgements (…) if and only if (a) the stereotype that is applied is accurate (…). (2021, 38)[9]

[8] Surprisingly, on page 56, the author replaces the verb "to lead" with the less demanding verb "to dispose".
[9] That is, if the stereotype leads to accurate judgements.

Moreover, Puddifoot's definition of the accuracy of a stereotype seemingly undermines her criticism of the dual factor view. According to Puddifoot, such view is flawed because it does not acknowledge a series of factors that may compromise the accuracy of the judgement resulting from stereotyping: the irrelevance of the stereotype to the particular decision problem, the misinterpretation, ignorance or selective recollection of the evidence, the discrediting of the testimony of the stereotyped person etc. If, however, a stereotype's accuracy is defined in terms of the accuracy of the resulting judgement, the factor of the stereotype's accuracy, which is central to the dual factor view, would already encompass all the epistemic factors that, according to Puddifoot, such a view mistakenly ignores (see 45-55). Indeed, these are all factors that may intervene between the endorsement of the stereotype and the ensuing judgement in each case; thus, factors that may compromise the accuracy of the judgement and, hence, according to Puddifoot's definition, also of the stereotype. If a stereotype's accuracy is defined in terms of the accuracy of the resulting judgement, then, requiring the former means requiring that the epistemic factors highlighted by Puddifoot do not materialise—or that, if they do materialise, they do not affect the judgement's accuracy.

Perhaps I am mistaken: Puddifoot never meant to define the accuracy of a stereotype in terms of the accuracy of the resulting judgement (there are indications to this effect in Chapter 2, Section 5). If so, an alternative interpretation must be given of the phrase "to respond in a way that is fitting with accurate statistical information", which is central to Puddifoot's definition on page 40. If this phrase does not mean "to judge accurately", what does it mean? A first possibility is that it means "to rely on accurate statistics in making one's judgement"—where the judgement may, however, turn out to be inaccurate. For a stereotype to lead the agent to rely on accurate statistics, one would expect that the stereotype itself must reflect such statistics. If so, though, why not defining the accuracy of a stereotype simply in terms of the accuracy of the generalisation it represents, rather than by reference to what it leads the agent to do? A second possibility is that the phrase means "to rely on accurate statistics in making one's judgement *and* to give to such statistics the appropriate weight". Again, the ensuing judgement may well be inaccurate, but for the stereotype to be accurate it must provide the agent with an accurate generalisation, and it must lead the agent to use this generalisation correctly. This means, for example, that the generalisation should not bring the agent to overlook case-specific evidence indicating that the case does not fit the generalisation, nor should it enable the recollection of only the evidence indicating that the case fits. It is unclear whether Puddifoot intended to define the accuracy of a stereotype according to this second hypothesis. It is worth noting, though, that this hypothesis presents the

problem discussed in the previous paragraph. If the accuracy of a stereotype is defined in terms of the appropriate weight given by the agent, in the particular decision problem, to the accurate generalisation reflected in the stereotype, then the dual factor view already includes some of the epistemic factors that, according to Puddifoot, are not part of it.

To conclude on Puddifoot's critical discussion of the single and dual factor views of stereotyping, Chapter 3 surprisingly lacks a definition of "relevance", notwithstanding that this notion is central to the discussion. Puddifoot argues that a flaw of these views is their failure to recognise that a factor that may undermine the accuracy of a judgement based on stereotypes is the possibility that the stereotype be triggered even when irrelevant to the judgement. To illustrate this point, she gives the following example of an irrelevant stereotype:

> A police officer approaches the car of a Black male, which has been pulled over for a minor traffic violation, e.g. one of his headlights is not working. The police officer asks the man to step out of the vehicle but he responds slowly and cautiously to the command. The police officer is offended at what he takes to be a threat to his authority. This triggers a stereotype associating the innocent man with crime; the police officer evaluates the man as a criminal and treats him with hostility; and this leads to an escalation of tension and hostility between the two individuals. The stereotype associating Black people with crime is triggered although the Black man has not committed a crime, only a minor traffic violation. (2021, 46)

Perhaps the omission of a definition of "relevance" is only remarkable for someone who, like me, works in the field of evidence law, where this notion is a cornerstone. But Puddifoot's example cannot be elucidated by falling back on the general understanding of this term as referring to the quality of being "related or useful to what is happening or being talked about".[10] In light of this understanding, I fail to see how the stereotype in the example, while admittedly spurious, is irrelevant to the decision-making of the officer. An officer's job includes preventing and detecting crime. True, the officer may have stopped the car for the sole purpose of fining the driver for a malfunctioning headlight. However, if, after stopping the car, the officer forms a suspicion that the driver has committed, is committing or will commit a crime, it is part of the officer's job to act on that suspicion. The stereotype being about the relationship between a social group and criminality, it is surely relevant to forming the

[10] See https://dictionary.cambridge.org/dictionary/english/relevance (accessed 9 June 2023).

suspicion; hence, to the decision-making of the officer in the case at hand. Again, the stereotype is spurious and may, therefore, lead the officer astray. But a charge of inaccuracy is not the same as a charge of irrelevance.

2. Egalitarian attitudes and epistemic pitfalls

On page 3, Puddifoot defines the concepts of "stereotype" and of "stereotyping" in comparative terms. Both involve "a social attitude that associates members of some social group *more strongly than others* with certain traits" (emphasis added). According to Puddifoot, what is required is not that the attitude reflects a stronger association than the actual one. After all, she makes clear that, according to her definition, stereotypes and stereotyping need not be inaccurate or lead to inaccurate judgements. The definition requires, instead, that the attitude reflects a stronger association for group members than for others. Therefore, an "egalitarian" attitude (that is, an attitude associating group and non-group members with equal strength with a certain trait) cannot be a stereotype. I am not convinced by the inclusion of this comparative definitional element: if, for example, vegans have a low carbon footprint and non-vegans have a higher footprint, isn't there a sense in which vegans are stereotyped if people associate both vegans and non-vegans with the carbon footprint that characterises the latter? Be that as it may, I won't dwell on the definition here.

In Chapter 4, Puddifoot discusses, among other things, the epistemic benefits of avoiding stereotyping. Consider an attitude that associates scientific expertise more strongly with men than women. It is a stereotype according to the above definition. Moreover, it is a stereotype that, to use Puddifoot's words, reflects an aspect of social reality in the United Kingdom: there are significantly more men than women in the sciences. This notwithstanding, Puddifoot argues that reliance on this stereotype involves significant epistemic pitfalls, introduced in Chapter 3. Conversely, an egalitarian attitude according to which men and women are equally likely to have scientific expertise does not track reality. And yet, Puddifoot argues that this epistemic cost is the lesser evil, compared with the epistemic pitfalls that would be avoided by endorsing such an attitude (ibid., 79). The bottom line is that, at least in some cases, avoiding stereotyping is preferable from an epistemic point of view, even if the stereotype tracks reality in some salient respect. In other words, in some cases endorsing an egalitarian attitude is preferable from an epistemic point of view, even if the attitude is inaccurate.

Puddifoot's conclusion may well be correct. Contrary to what she suggests, though, the egalitarian nature of the alternative attitude cannot alone guarantee the less costly epistemic outcome. The author writes that

> The possession of the attitude that women and men are equally likely to have scientific expertise will guard against various tendencies that accompany the stereotype associating scientific expertise more strongly with men than women. (2021, 75)

The reader will remember from Chapter 3 that these tendencies, or pitfalls, include misinterpreting, ignoring, or remembering selectively the evidence, so that the recognised evidence fits and confirms the stereotype, as well as giving a credibility deficit to the testimony of those who are stereotyped. Now, if I endorsed an attitude according to which both men and women are *extremely unlikely* to have scientific expertise, I would probably fall prey to (at least some of) these pitfalls when assessing the testimony of a woman (or of a man) scientist—e.g., the base rate that I endorse may lead me to give this testimony a credibility deficit. This is notwithstanding that my attitude is egalitarian (hence, not a stereotype). The egalitarian nature of the attitude may mean that the distribution of epistemic errors is also egalitarian: other things being equal, errors will equally affect the men and the women whose expertise I judge. Pace Puddifoot, though, it does not necessarily mean that the above epistemic "tendencies" are avoided, or that they are less pronounced than in the case of stereotyping.

3. The moral encroachment approach to stereotyping

Chapter 5 analyses the relationship between epistemic and ethical demands in the case of stereotyping. Puddifoot argues that this relationship is a complex one: depending on the context, epistemic and ethical demands may align or may clash. Here is a brief illustration of her account. Reliance on a stereotype that reflects an accurate generalisation (say, associating Black people more strongly than non-Black people with a certain medical condition) may lead to an accurate judgement (say, an accurate medical diagnosis of a Black patient). But it may also trigger further stereotypes (say, that associating Black patients more strongly than non-Black patients with uncooperativeness), which may foster a poor interaction between the individuals involved (perhaps with detrimental medical consequences for the Black patient). This is a case in which epistemic and ethical demands appear to clash: only the former justify stereotyping. In earlier chapters, though, Puddifoot has argued that stereotypes (including those that reflect accurate generalisations) may lead to significant epistemic pitfalls, such as the misinterpretation, ignorance or selective recollection of evidence, and

the unwarranted discounting of the testimony of the stereotyped person. Hence, reliance on a stereotype may lead to a series of inaccuracies (say, unwarranted distrust in the patient's description of their symptoms). In this case, epistemic and ethical demands may align: neither justifies stereotyping.

Puddifoot contends that the moral encroachment approach to stereotyping cannot capture the complexity of the relationship between epistemic and ethical demands. Her conclusion, though, seems too quick. To see why, let's start with Puddifoot's brief illustration of the approach. According to its defenders

> [E]thical and epistemic demands do not conflict because moral considerations determine whether it is epistemically permissible to engage in stereotyping (…). Where there are high moral stakes in a situation in which a judgement is made, high evidentiary standards need to be met in order for a judgement to be justified or rational or to constitute knowledge. In cases where people might engage in stereotyping, there will often be high moral stakes, and these stakes will raise the evidentiary standards. Those who engage in stereotyping will not meet the high evidentiary standards. Therefore, stereotyping will not be justified or rational. (2021, 114)

Puddifoot highlights two main problems with this approach. First, it is "overly simplistic" in its failure to acknowledge that epistemic and moral considerations may, indeed, clash. Second, it mistakenly assumes that stereotyping is never justified in high-stakes situations: "sometimes high-stakes situations demand stereotyping, because stereotyping can be an efficient way to achieve both ethical and epistemic goals, like correct diagnoses and treatment decisions" (ibid, 115). I am not particularly interested in engaging with the first criticism. The question as to which are the possible arrangements between the epistemic and the ethical "vectors" strikes me as more formal than substantial. What matters most in assessing an approach to stereotyping is whether the approach offers appropriate guidance to those who may stereotype. In this regard, I agree with Puddifoot that it is mistaken to foreclose the possibility of stereotyping in high-stakes situations. However, I believe that a moral encroachment approach to stereotyping can account for cases in which stereotyping is (epistemically and ethically) justified, notwithstanding the high stakes involved. In order to defend this thesis, though, I will first defend the equivalent thesis formulated with reference to the distinct theory of pragmatic encroachment.

The essential tenet of pragmatic encroachment is that whether someone knows *that p* (say, that a patient has a given condition) depends on the importance that getting the matter right has for this individual, thus on their practical interests involved. It is possible to operationalise the pragmatic encroachment approach by modelling a decision problem using decision theory. This will allow us to account for the stakes, as well as to identify the evidentiary standard (or probability threshold) that must be satisfied for someone to be justified in acting as if the proposition at issue were true, given the stakes. Under pragmatic encroachment, being justified in acting this way means having an "outright" belief in the proposition—a belief necessary for knowledge. A full illustration of this model is beyond the scope of this article.[11] It suffices to point out the following. The model requires identifying the possible outcomes of the decision problem (say, the problem whether to diagnose a condition). These are: the correct outcomes consisting in acting as if the proposition were true when it is, indeed, true (say, diagnosing a condition when the patient has such a condition) and in acting as if the proposition were false when it is, indeed, false (say, not diagnosing a condition when the patient does not have such a condition); and the mistaken outcomes consisting in acting as if the proposition were true when it is, indeed, false (say, diagnosing a condition when the patient does not have such a condition); and in acting as if the proposition were false, when it is, indeed, true (say, not diagnosing a condition when the patient has such a condition).

Once the possible outcomes are identified, a value should be assigned to each outcome, reflecting the preferences, or practical interests, of the decision-maker. The model then allows to identify a probability threshold the satisfaction of which justifies acting as if the proposition at issue were true, given such values (i.e., the stakes). For the purposes of the pragmatic encroachment approach, this probability threshold corresponds to the attitude of outright belief: the decision-maker outright believes the proposition at issue if and only if, their degree of belief in that proposition corresponds to a probability equal to, or greater than, such threshold. Therefore, outright belief in the proposition at issue implies that the agent is justified in acting as if that proposition were true.

Notably, the model does not foreclose the possibility that the decision maker has an outright belief in a proposition notwithstanding that their degree of belief in that proposition is fairly low *and* the decision problem hinging on the truth of that proposition involves high stakes. It is possible

[11] In essence, however, a decision problem hinging on the probability of a particular state of affairs (e.g., whether a patient has a given condition) can be modelled similarly to the decision problem of criminal adjudication, which hinges on the probability of the defendant's guilt. For a decision-theoretic approach to the latter, see Picinali (2022, chs. 3 and 4)

that acting as if the proposition were false when it is, indeed, true, is an extremely costly outcome, whereas acting as if the proposition were true when it is false, presents moderate or little cost. For example, not diagnosing a condition when the patient has such a condition may quickly lead to the patient's death, whereas diagnosing the condition when the patient does not have it may result in subjecting the patient to a treatment with some beneficial consequences and few side effects. Depending on the values of the correct outcomes, in this situation the model may well indicate that the evidentiary threshold for outright belief in the proposition at issue is low: the agent may outright believe and, hence, know that the patient has the condition (and may be justified in acting accordingly, that is, in diagnosing the condition and giving treatment) even in the presence of weak evidence that the patient, indeed, has it. This is a high-stakes situation (one of the mistaken outcomes, the false negative, is very costly) in which stereotyping may well be warranted in accordance with the essential tenet of pragmatic encroachment, notwithstanding that the stereotype linking members of a group to which the patient belongs with the condition (more strongly than non-members) may not provide robust epistemic support.

The essential tenet of moral encroachment is that whether someone knows *that p* depends on the moral stakes of the decision problem. The practical interests of the decision maker, with which pragmatic encroachment is concerned, may not track the moral stakes (e.g., the agent may care very little about whether the patient will suffer harm in case of a false negative, being chiefly concerned with the costs of treatment for the hospital's coffers if the condition is indeed diagnosed, and with the repercussion these may have on a desired salary increase). Therefore, pragmatic encroachment and moral encroachment may lead to different conclusions as to whether the agent knows *that p*.

In standard decision theory the value function reflects the decision maker's preferences, which is why pragmatic encroachment lends itself to modelling through decision theory. However, one can construct the value function as tracking moral value, that is, as reflecting the preferences of the morally conscientious agent.[12] Once this condition is added, the argument can be rerun with reference to moral encroachment. In essence, the point is that there will be situations in which the false negative has such a high moral cost that the threshold for outright belief (and, hence, for knowledge) will be relatively low; sufficiently low to be satisfied by stereotyping. In the

[12] One may object that decision theory cannot capture essential aspects of a deontological theory, such that a deontologist cannot model moral problems with decision theory. On this issue see the work of Lazar (2017).

medical example in which a false negative (missed diagnosis) leads to a quick death and a false positive (false diagnosis) leads to some health benefits and mild side effects, the doctor may well be morally warranted to follow a stereotype linking members of a group to which the patient belongs with the medical condition at issue (more strongly than non-members).

4. Evaluative dispositionalism

On page 164, Puddifoot introduces her theory of "evaluative dispositionalism" (ED) as a theory to assess the justification for holding "stereotyping beliefs", that is, "beliefs that encode generalizations about social groups, associating all group members more strongly than non-group members with some feature" (2021, 145). This theory is introduced after arguing that existing accounts of epistemic appraisal (that is, upstream, downstream,[13] and static accounts) cannot capture some of the epistemic faults associated with holding stereotyping beliefs. According to ED

> [A] complete epistemic evaluation of an act of believing should focus on both (a) the dispositions that are displayed in believing, and (b) the dispositions that are possessed due to believing. (2021, 164).

Effectively, ED is a combination of the existing accounts.

A preliminary issue with Puddifoot's analysis is that it fluctuates between presenting ED as a theory for the appraisal of beliefs tout court and as a theory for the appraisal of stereotyping beliefs only. On pages 158 and 188, Puddifoot states that she is only concerned with the more modest task of offering a theory that targets stereotyping beliefs. And yet, the formulation of ED that I have just reproduced is couched in general terms. What is more puzzling is that, in illustrating the theory, Puddifoot relies on examples of beliefs that are not stereotyping beliefs (e.g., the belief that Manchester City is going to win the Premier League). Perhaps this fluctuation betrays greater (and perfectly sensible) ambitions than those declared. Be that as it may, ED is not an entirely convincing theory even when circumscribed to stereotyping beliefs; or, at least, Puddifoot does not show that it is.

On pages, 181-2 Puddifoot offers a hypothetical case in which two individuals harbour the same stereotyping belief and yet, according to her, this belief deserves different epistemic evaluation in the two cases (it is

[13] To be sure, downstream accounts are, by and large, a creation of Puddifoot (see 2021, 141-143).

justified in one case only, perhaps) because only one individual has dispositions to make epistemic mistakes due to holding the stereotyping belief. Here is the hypothetical case.

> Consider two people. Nora is a female scientist who has 30 years of experience. She is a feminist and as a result pays close attention to the representation of women in the sciences. She notices over time that a gender gap in the sciences never goes away: there are consistently more men than women in sciences, and therefore consistently more men than women with scientific expertise. Nora therefore harbours a stereotype associating men more strongly than women with scientific expertise. She harbours and endorses the social attitude *men are more likely than women to have scientific expertise*. However, the stereotype does not distort Nora's judgements of individual women scientists and their levels of expertise. The stereotype does not make her assume that women scientists are more similar to each other than they really are, or that they are less similar to men scientific experts than they really are. She does not misremember the features of women scientists due to the operation of the stereotype. And so on. Instead, Nora judges women scientists on the basis of the skills, expertise, and potential that they display in their work, with the stereotype only operating to allow her to understand the challenges that they are likely to have faced as a minority in the profession.
>
> Ned is also a scientist with 30 years of experience. He is not a feminist. He has also registered that women are underrepresented in the sciences and therefore harbours a stereotype associating men more strongly than women with scientific expertise. He endorses the social attitude *men are more likely than women to have scientific expertise* and thereby harbours the same stereotype as Nora. However, the stereotype that he harbours permeates his thought, influencing all of the judgements that he makes about individual women scientists, and about the relative merits of men and women scientists. He makes errors such as misremembering the attributes of his women colleagues, misinterpreting ambiguous behaviours as indicating a lack of expertise, assuming women colleagues are more similar than they really are, and so on.
>
> These two characters harbour the same stereotype: that men are more likely than women to have scientific expertise. Their stereotypes are formed on the basis of the same evidence:

> evidence about the underrepresentation of women in the sciences. But on the evaluative dispositionalist account the characters and their act of believing deserve different epistemic evaluations because the characters differ in the dispositions that they have due to possessing the stereotyping belief: Ned has dispositions to respond poorly to the evidence while Nora does not. This seems to be precisely the right result.

Both Nora and Ned hold the stereotyping belief that men are more likely than women to have scientific expertise. As it happens, this belief is accurate in Nora's and Ned's society (remember that, under Puddifoot's definition, stereotypes need not be inaccurate). Since they both hold this belief, one would expect that both will fall prey to a series of epistemic pitfalls which, as argued by Puddifoot in earlier chapters, are produced by harbouring stereotypes. And yet, this is not the case: only Ned incurs the pitfalls. Puddifoot suggests that this is because Nora is a feminist, whereas Ned is not. Consequently, while holding the stereotyping belief, Nora will not, e.g., misinterpret evidence concerning the scientific expertise of a woman or give a credibility deficit to the testimony of a female scientist. In this scenario—Puddifoot concludes—ED indicates that Nora is justified in holding the stereotyping belief, whereas Ned is not.

What is perplexing about this conclusion is that, as accepted by Puddifoot, the epistemic pitfalls which Ned incurs are not, or not just, "due to" his holding the stereotyping belief. If they were just due to Ned's holding the belief, one would reasonably expect Nora to incur such pitfalls as well. Perhaps the fact that Ned holds the stereotyping belief is indeed irrelevant to the occurrence of the pitfalls, these being entirely brought about by pre-existing dispositions such as the rejection of feminism. Perhaps holding the belief is not irrelevant after all: it acts as an enabler or as an enhancer of pre-existing dispositions that bring the agent to incur the epistemic pitfalls. Now, if holding the stereotyping belief is irrelevant, then there is no reason to accept that the pitfalls should be factored into the assessment of the justification of holding such belief. In fact, ED says that they shouldn't, since they are not "possessed due to" holding the belief. If, instead, holding the stereotyping belief enables or enhances pre-existing conditions, then there is an explanatory story to be told and evaluated, the complexity of which is not accounted for in the current formulation of ED.

I suggest that ED would be improved by clarifying how significant should be the role of the belief in bringing about the pitfalls for the pitfalls to be factored into the assessment of the justification of holding the belief. In the current formulation of ED, any causal (or enhancing) role of the belief seems sufficient. However, this may be too strict a position to take. After

all, it is doubtful whether it would be epistemically unjustified to hold a stereotyping belief that is accurate (as that in the example) and contributes only minimally to bringing about the epistemic pitfalls. It is also possible that the question of the role played by the stereotyping belief in the aetiology of the pitfalls is intractable: there is simply no way of ascertaining the nature and extent of this role in any given case. If so, ED is in even greater trouble.

5. Conclusion

This article was an attempt to highlight some flaws in Kathy Puddifoot's book. First, I criticised her treatment of the "single factor view" and of the "dual factor view" of stereotyping, raising questions about the examples and the conceptual apparatus that Puddifoot relies on. Second, and contra Puddifoot, I argued that endorsing an egalitarian attitude may not avoid the epistemic pitfalls associated with stereotyping. Third, I argued against Puddifoot's claim that a moral encroachment approach cannot justify resorting to stereotypes in high-stakes situations. Fourth, and finally, I argued in favour of enhancing Puddifoot's "evaluative dispositionalism" with a clarification of the causal role of the stereotyping belief vis-à-vis downstream dispositions.

These flaws notwithstanding, the book is a fascinating and engaging read, highly recommended to epistemologists and legal scholars alike.

Acknowledgments

I am grateful to Kathy Puddifoot and Federico José Arena for their comments on an earlier draft.

REFERENCES

Lazar, Seth. 2017. "Deontological Decision Theory and Agent-Centered Options." *Ethics* 127 (3): 579–609. https://doi.org/10.1086/690069.

Picinali, Federico. 2022. *Justice In-between: A Study of Intermediate Criminal Verdicts*. Oxford and New York: Oxford University Press. https://doi.org/10.1093/oso/9780198864592.001.0001.

Puddifoot, Katherine. 2021. *How Stereotypes Deceive Us*. Oxford: Oxford University Press.