



UNIVERSITAT DE
BARCELONA

Faculty of Philosophy

Master in Cognitive Science and Language

Academic year 2022/2023

***Moral Perception: seeing the Good, the Bad and
the Ugly***

Author: Lucía González Arias

Supervisor: Josefa Toribio Mateas

June 2023

Table of Contents

Acknowledgements	3
Abstract.....	4
Introduction.....	5
Revisiting AMP: the empirical side.....	6
Attentional moral perception, or methodological artifacts?	6
The moral pop out effect.....	7
Binocular rivalry	10
Attentional Patterns in Moral Dilemmas	12
Gaze and cheating	14
Gaze and generosity	15
Asymmetries: what does data <i>really</i> say?.....	16
Revisiting AMP: the theoretical side.....	17
Challenging the priors.....	18
The moral and the social	20
Neural substrate and clinical cases	27
Conclusions.....	30
References.....	31

Acknowledgements

I would like to express my most sincere gratitude to my advisor, Pepa, for her invaluable guidance, corrections, and suggestions. And, above all, for opening the doors for me to a whole new world I always longed to be a part of.

I also want to thank my mother, for always pretending to listen to me while I ramble about these topics, and still love me unconditionally. My sister, too, for all the spells she put on me to help get to where I am now. And to my dad, too, who I am sure would have been part of my hostage-audience as well.

To my lovely friends from the J.O.T, thank you for being always there.

Of course, Patricia, my cats, the bumblebee that I befriended, and the birds that sing with me in the balcony, thank you for being my sources of sanity throughout this journey. You have made me reconsider my career path as a full-time Disney princess.

Lastly, this is also for you, Medina. I like to imagine that you now reside within every page that I write, in which I devote myself to those topics I “rebelliously” introduced in all the papers I had to hand you in. “*Your work is full of irrelevances*”, you once told me, “*But, oh God, such interesting ones. Do not dare to stop doing that*”. Well, look at me now!

Abstract

Attentional moral perception (AMP) is the most recent and less-controversial version of moral perceptualism, the claim that we can represent moral properties such as being right or wrong in perception. AMP features an attentional, albeit perceptual sensitivity to moral difference-makers that need not involve the perceptual representation of the moral properties themselves. In this essay, I examine the empirical and theoretical basis of this proposal, delving into the methodological imprecisions and problematic implications that comes with endorsing it. On the empirical side, I cast doubt on the validity of the data and the fitness of the experimental designs typically invoked to defend the view. On the theoretical side, I articulate a more controlled example and, borrowing a classic argument from Speaks (2010), I argue that AMP ultimately needs to commit to phenomenologically rich ‘attentional states’ and renounce the transparency of perception. Clinical and evolutive aspects of AMP are also discussed.

Key words: attentional moral perceptualism; moral perceptualism; moral cognition.

Introduction

Our daily life is plagued with morally charged situations, as it is our language with expressions to describe them. This comes just “naturally”; we do not need much thinking to label an action as morally wrong or right. Instead, it is almost intuitive to plainly *see* them as such. Seeing the good in someone, the bad of some situation, are quite common expressions to characterize our experience of morally charged situations. Of course, this is figurative language—or is it grounded in something else?

Moral perceptualism has been an area of active philosophical inquiry for many years. Recently, the debate seems to have been stirred up, as a significant corpus of empirical and theoretical arguments has arisen to account for our moral properties’ recognition abilities.

The defenders of moral perceptualism (MP) are generally committed to the idea that we can literally see (at least some) moral properties, such as the rightness or wrongness of an action. This means that some moral properties constitute a part of the content of our perceptual experience (Werner, 2020). In fact, following Vance & Werner (2022), this description would account for what they call Contentful Moral Perceptualism (CMP). CMP has been heavily criticized both in the theoretical and empirical field (see, e.g., Väyrynen, 2018; Cowan, 2015). It has been argued, for instance, that, while perception must engage the perceiver in causal connections with the perceived object, moral properties are causally inert, so they cannot be represented in perception (see the causal objection formalised in McBrayer, 2010). The perception of moral properties, too, has been argued against by claiming that there is not a typical look of a ‘moral property’ to be subject to perceptual representation (Reiland, 2021). Moreover, moral beliefs, as a top-down influence, have been also proposed to yield explanatory power as an instance of cognitive penetration, so perception of moral properties would not really add anything to the picture (see the Redundancy argument by Reiland, 2021).

Nonetheless, there is a type of MP that, *prima facie*, dodges CMP's critics and provides a good advocate for MP. This is the case of what Vance & Werner call Attentional Moral Perceptualism (AMP). AMP's main feature is that there is no commitment to the representation of moral properties themselves; instead, AMP features the role of perceptual, *attentional* mechanisms that are sensitive to moral difference-makers, cues that indicate that a situation is morally charged. Vance & Werner (p.9) provide the following illustrative example, regarding a car running over a pedestrian after skipping a red light: "*the traffic lights being red when the car hit the pedestrian is morally relevant in that context, but in most contexts in our social environment, something being red is not a morally relevant feature*". This sensitivity is said to be reflected in attentional patterns in the perceptual experience, that ultimately can influence moral cognition.

This is not a trivial stance. Committing to the naturalisation of morality in the attentional realm, as I will elaborate in this essay, has great consequences in the philosophical, psychological, and social domain.

Revisiting AMP: the empirical side

Attentional moral perception, or methodological artifacts?

Empirical data is a powerful support for a theory in philosophy of mind and perception. Accordingly, it is standard to invoke psychological data to support such theories. Nevertheless, jumping into the "empirical pool" comes with some risks. Of course, not all empirical evidence is valid: psychologists have been guilty of appealing to some "methodological tricks" to favour certain interpretations of their own results.

In this section, I will revise the central empirical arguments in favour of the AMP. To do so, I will take as a reference Vance & Werner (2022)'s work, which gathers the most relevant empirical studies and phenomena that MP theories (AMP, in particular) rely on.

Additionally, I too will examine the general conclusions that can be drawn from all the data, highlighting the findings' asymmetric nature, a possible rejoinder, and, overall, the problem of designing an appropriate experiment to test AMP.

The moral pop out effect

Methodological constraints

The moral pop-out effect is defined by Gantman and Van Bavel (2014, p. 23) as the phenomenon by which “*perceptually ambiguous moral stimuli are more likely to reach perceptual awareness than matched non-moral stimuli*”. To prove its existence, they conducted a series of studies based on the lexical decision task (LDT) paradigm.

In experiment 1, they presented participants with moral and non-moral words, as well as pseudowords, and measured reaction times in telling if the stimuli comprised real words. In experiment 2, they tested whether the moral pop-out effect was only evident for perceptually ambiguous stimuli. In experiment 3, the effect of harm as a prime for the moral pop-out effect was tested. Overall, the results pointed out that moral words were responded to more quickly than non-moral words and, furthermore, that this effect was independent to valence or arousal's own effect, and was not mediated, either, by exposure to harm. Last, they suggested that, in reaching conscious awareness, perceptually ambiguous moral stimuli may require fewer prerequisites than non-moral stimuli.

At first glance, these results support the evidence of a moral pop out effect. Still, one might want to take a careful look at what seems to be a methodological artifact in experiment 1; according to the authors, “*there is no effect of word type on reaction time when we regress log-transformed reaction time on word type ($p = .90$, $z = .14$)*” (Gantman & Van Bavel, 2014, p. 24). According to this footnote, they did not assess the reaction time (RT), due to the short duration of the presented stimuli. Although it is possible that just one of the main measures in psycholinguistics, Reaction Times and Accuracy, is affected by the word type condition, it

would be sensible to set up an experiment to see whether there is actual conflicting data. This is, while Accuracy is a perfectly valid measurement, RTs are often more informative for cognitive and perceptual fine-grained analyses, since it is a more robust measure for attention than Accuracy (Prinzmeta et al., 2005). Nevertheless, the important point revolves around the possible incoherence between the results: if Accuracy can account for the studied phenomenon, it follows that a preference for morally charged words should be seen in the RTs measure as well, as they are dependent variables.

Another important methodological consideration when evaluating Gantman and Van Bavel's findings is, of course, their sample size. Fairly enough, the authors themselves acknowledged that it (n=19) may have been inadequate for the purposes of the study: "*Though we had previously decided to determine our sample size by the end of the semester, reviewers suggested that we rerun the study for an additional semester to increase statistical power*" (p. 5). Note that, unless it is a one-case-study, experiments typically require a minimum sample size of 30 participants to obtain statistically significant results (Chang et al., 2006; Cohen, 1988). Therefore, the small sample size in this study may compromise generalizability of its findings to other populations or contexts, i.e., the findings might not be replicated outside the context of that specific experiment.

The sample can be threatening generalizability due to its homogeneity, too. The participants in the study were exclusively university students, a group that may not accurately represent the wider population regarding an important variable: the stage of moral development. Although constructivist approaches are decaying, it is still useful to refer to Kohlberg's (1963) theory of moral development, or Gilligan's (1977; 1982), to address how moral reasoning and moral sensitivity changes over development¹. In this regard, the sample

¹Note that both the development of moral sensitivity and moral reasoning, under AMP lenses, would stem from the improved ability of perceiving moral-difference makers, as a prerequisite for both

may not be representative of the general population: college students are assumed to undergo common experiences that significantly mark their moral development (Rest & Narvaez, 2014) and makes it different from the one of non-students.

In addition, there is a psychometric construct worth examining that has been systematically neglected in the experimental field over the past decades: ecological validity. Roughly, this type of validity tries to capture how accurately laboratory conditions resemble real-life situations and whether the results obtained in the former environment can be generalized to the latter (Keidser et al., 2020). Regarding our topic, while laboratory-based studies offer many advantages (other types of validity, reliability, an enhanced control of strange variables...), the context in which moral decisions and attentional moral ‘perception’ take place in real life is much more complex and multifaceted (Greene & Haidt, 2002). Therefore, it is possible that the study's findings may not fully reflect how moral perception operates in natural settings, when we are “rationally bounded” (Simon, 1990), but rather how it operates under particular, artificial, conditions.

What is more, it is difficult to adduce a “moral pop-out effect” relying solely on a task in which participants are reading *words*. Natural environments in which AMP is supposed to take place are not prompted by seeing the written word “MURDER” in neon lights in the street, but rather by witnessing complex, often ambiguous, morally charged situations and behaviours. Definitely, ecological validity is going to constrain, as we will see in the next sections, much of the optimistic empirical support to AMP. This is why I strongly advocate and urge the revision of these studies’ conclusions.

Last, I want to stress the importance of controlling as many psycholinguistic variables as possible when designing a moral LDT, to make sure that the effect cannot be better explained

instances would be to be actually capable of detecting those situations, hence becoming V&W’s ‘virtuous agents’).

by them, rather than by the moral charge of the word. In this regard, Firestone & Scholl (2016) pointed out how many stimuli pools that are used in paradigmatic cases of “moral pop-out” experiments are biased, as they are not matched in every relevant psycholinguistic variable. In Gantman & Van Bavel’s case, frequency was neglected.

Alternative explanations: semantic priming

Aiming for the most parsimonious and coherent explanation, one alternative to invoking moral perception as an underlying mechanism for the observed data is to rely on the phenomenon of semantic priming (Toribio, 2023). It consists of the activation of a semantic network due to the previous presentation of semantically related stimuli, which can deteriorate or improve performance on the task. Indeed, this view has been endorsed by Firestone & Scholl while arguing against Gantman & Van Bavel. By replicating the latter’s experiment with fashion and transportation categories instead of moral ones, Firestone & Scholl found that the same pop-out effect could be elicited with just a general ‘category-belonging-elements’ vs ‘category-non-belonging-elements’ design. This further supports the semantic priming hypothesis: Gantman’s & Van Bavel’s results, as Firestone & Scholl (2015, p. 412) say, “*should be interpreted in terms of ‘back-end’ memory retrieval rather than ‘front-end’ visual processing*”. Visual awareness has apparently no role here; instead, the presented words in both experiments lowered the standard threshold for activation of the semantically related words.

Binocular rivalry

In a binocular rivalry experiment, participants are presented with different images in each eye at a time, with little to no control over which image dominates the experience. In Anderson et al. (2011) experiment, participants were first shown pictures of neutral faces paired with positive, negative or neutral social actions. Then, they conducted a binocular rivalry task with the same pictures of neutral faces and unrelated stimuli. They reported that faces previously associated with socially negative actions were more likely to dominate conscious

awareness compared to faces associated with socially positive or neutral actions. If one image is preferentially attended, they hypothesise, it can dominate experience—that is, it reaches conscious awareness more frequently and for a longer time. They conclude that “*faces previously paired with socially negative actions tended to dominate experience, indicating preferential attention to them over faces paired with socially positive and neutral actions*” (p. 11). Vance & Werner (p. 12) expand the discussion and reason that “*it is likely that bad agents’ faces dominate experience over good and neutral agents’ faces, because attentional mechanisms are sensitive to the moral status of agents’ actions and prioritize bad agents’*”.

Now, it is important to take into account that “socially positive”, “socially neutral” and “socially negative” actions, all refer to violations of normative, standard social norms. Fireston & Scholl (2016) regarding Anderson et al.’s findings, pointed out that (1) the apparent moral pop-out effect worked with actions that were socially negative, but not necessarily immoral (e.g. being in a car accident while borrowing your friend’s car), and (2) failed to show up regarding morally good actions (e.g. helping the elderly). In addition to this, Anderson et al. mentioned (although they did not go into much detail) that neither socially positive nor neutral actions were preferentially attended compared to negative actions... but the moral pop-out effect is meant to be orthogonal to the overall valence of the input! It seems, thus, that morality per se does not seem to be driving the reported effect.

Instead, what binocular rivalry might be revealing can be just the “leftovers” of a common bias (Toribio, 2023). The negativity bias can be defined as the tendency for negative information to have a stronger and more lasting impact on our cognitive and emotional processing than positive information (Rozin & Royzman, 2001). As many other cognitive biases, it can influence our attention, memory, and decision-making processes. What’s more, in terms of neurobiology, studies have shown that the presentation of negative, compared to neutral and positive stimuli, prompts higher activity in the amygdala (Gamer et al., 2010;

Cunningham & Brosch, 2012). Note that this brain region is involved in processing emotional values and detecting potential threats in a very automatic fashion (Phelps, 2006), and this is directly linked to the Social Learning Theory (SLT; Bandura, 1977): individuals easily learn to avoid those who engage in socially negative actions, having observed the negative outcomes associated with such behaviour. Logically, the tendency to encode socially negative actions as threats is an adaptative response that is instantiated in a bias in, among other processes, attention. And, even if the social value of the conditions was controlled, attention being a part of the preferential coding of negative information does not entail that this role is embedded in the perceptual experience.

In sum, given the “natural” salience of negative information and its importance for survival and threat detection, it is safe to cast doubts on whether (1) this experiment is just showing that our attention is preferentially driven to negative social cues rather than negative moral cues and (2) this experiment add something to AMP’s pretensions of attention as a core perceptual mechanism. Additionally, as I have previously mentioned, in absence of pairwise comparison analyses, the presumed effect would only work for the morally/socially negative condition, not the positive one.

Vance and Werner also rely on some studies about attentional patterns in moral dilemmas, gaze and cheating, and gaze and generosity, all of which allegedly support AMP. I’ll discuss these separately below.

Attentional Patterns in Moral Dilemmas

In Foot’s (1967) study, some peculiar attentional patterns are described in the context of a variation of the trolley moral dilemma, in which a person A or person B had to be sacrificed. Generally, participants attended more to the person whom they would eventually sacrifice. This, according to the authors, is most likely reflecting empathetic concern for victims or a need to

process victims' traits as part of the decision-making process. Participants generally avoided attending to the victim whom they had sacrificed, possibly reducing cognitive dissonance between feelings of guilt about and endorsement of sacrificing that person. This, however, does not seem to provide a clear support for AMP. Surely, difficult decisions prompt a more intense reflection, which inevitably comes with enhanced attention to all possible information. Attentional patterns also change due to cognitive and emotional influences or states. These seem more instances of cognitive penetration, top-down influences, and consequent attentional-patterns adjustments, rather to AMP. The lack of automaticity in decision-making paradigms, although it is not a knock-down condition (see Wu, 2011) surely constrains the approach to attentional processes that one must commit to in order to defend AMP.

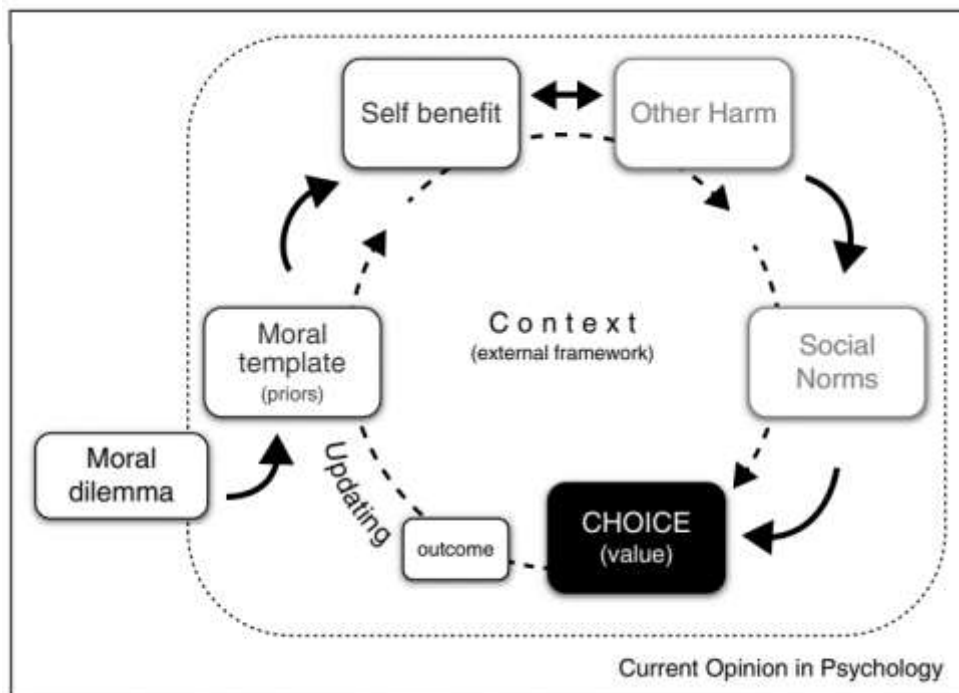
Regarding Garon et al.'s (2018) results, yet again, nothing really seems to stem here in support of AMP. The number of fixations on morally "interesting" features just exemplifies the former case: in facing a moral dilemma, or any complex decision, in general, we seek information to commit to a judgment. The leap from attentional patterns being an important factor in the encoding of information in a bounded-rationality framework, to attentional patterns being a central feature of the perceptual experience is unclear to me.

Additionally, ecological validity is a particularly relevant concern in the assessment of moral dilemmas (Kneer & Hannikainen, 2020). There is a huge issue with external and mutual internal validity in these experimental settings, since the laboratory conditions in which these decisions take place can in fact alter the results. For instance, one main caveat is the lack of real consequences of the moral dilemma. Feeling observed is also a huge noisy variable, which contributes to the biasing of the responses due to social desirability.

Apart from that, the hypothetical scenarios used to study moral decision making typically ignore the influence of socio-emotional factors and contextual tensions endemic to real moral dilemmas (fig. 1).

Figure 1

A dynamic model of moral cognition (Van Bavel et al., 2015)



Gaze and cheating

Hochman et al. (2016) set an experiment in which participants had to respond on which side of a divided screen there were more dots, and they would get an incentive depending on their performance. At first, what was more valued was being accurate, but, eventually, the incentive changed and benefited cheaters (systematically choosing one of the sides, independent of the number of dots). Critically, throughout eye-tracking, they found that when individuals cheated, they paid less attention to the relevant information that could expose their cheating (such as missing dots), unlike when they responded accurately in either incentive condition.

Contra Vance and Werner's interpretation of this study as support for AMP, directing attention towards self-serving aspects of the situation and avoiding cues of wrongdoing when cheating occurs is easily an instance of the confirmation bias, by which information seeking and coding is suited to our beliefs and expectations (Peters, 2022), being one of our beliefs that, for instance, we are good people that do not engage in cheating. This is, actually, a modality of the confirmation bias that can be defined as the self-serving bias (Mezulis et al., 2004).

As previously mentioned, the results point towards the effect of a cognitive bias, which prompts changes in the allocation of attention, rather than changes in the perceptual experience itself.

Gaze and generosity

The Dictator Game (DG) is a common task to measure prosocial behaviour, and, in this case, it has also been used to assess how perceptual attention patterns can predict the level of generosity in strategic games. The task goes like this: experimenters present subjects with proposals consisting of monetary trade-offs between themselves and an anonymous partner. They have to either accept or reject proposals in each trial, evaluating their gains and losses. Vance & Werner highlight Teoh et al. (2020)'s and Rahal et al. (2020)'s results on different DG experimental settings. Overall, they found that participants' generosity was affected by (1) under time pressure, whether they fixate their attention on self-relevant or other-relevant information, (2) whether they are dealing with ingroup or outgroup members. Together, results are taken by Vance & Werner (2022, p.17), as indirectly showing that "*patterns of attention can differentially track morally relevant features, and that attending to morally relevant features (and away from irrelevant distractors) can make a difference to moral decision-making*".

The correlation between attentional patterns in vision and participant's level of generosity, however, is not something that aligns exclusively with AMP's predictions. As in

the other sections' cases, these results are a showcase of what is it like to make a decision under bounded-rationality terms. In their particular setting, Rahal et al.'s results just profile the in-group bias, which causes people to preferentially attend to congeners' information. As for Teoh et al.'s experiment, the authors clarify that these findings could not be replicated if there was no time pressure (a bounded-rationality setting!). In these cases, heuristics and cognitive shortcuts are the ones mainly driving the decision-making process (Gigerenzer, 2004), which seems to put aside perceptual attention's role in favour of the influence of cognition on lower-level processes.

Asymmetries: what does data *really* say?

Apart from the methodological caveats previously mentioned, there is a crucial point in AMP's formulation that does not quite fit both the presented and further data. A theory that claims that we, as moral agents, are somehow sensible to moral properties, must account for both ends of the moral spectrum. Binocular rivalry studies, as previously explained, demonstrate how we are attuned to *negative* stimuli. Even if we take the results as proof of some kind of "attunement to moral-cues", the conclusion will still go only one way, and that is that we are attuned only to morally *negative* stimuli². As for the experiments that claim to prove otherwise, the problem of asymmetry still holds. In LDTs, there are no pairwise comparisons that allow us to see the direction of the difference between morally charged and morally neutral words—the main conclusion that can be drawn from these studies is that morally charged words are more attended than morally neutral words. Still, this is not just a matter of inadequate data analyses, but also methodological imprecisions: if morality cannot be reduced to psycholinguistic properties such as valence or arousal, there must also be an extra axis to a

² For the purpose of integrating external studies on the topic, AMP might rely on the "positivity bias" as a last resource to defend the bidirectionality of attentional moral perceptualism. Still, the positivity bias has only been tested under the same methodological constraints that the addressed studies suffer from: lack of control group, indirectness of the measurement, uncontrolled variables...etc (see, e.g., Mezulis et al., 2004; Peeters, 1971; Unkelbach et al., 2020).

word's definition: its moral valence (Shuman et al., 2013). This way, to assess AMP with a LDT, we would have to control these word conditions: negative/positive, high-arousal/low-arousal, morally-charged/morally-free, AND morally-positive-valenced/morally-negatively-valenced. As an example, euthanasia might be a plausible candidate that yields negative valence, as it entails death, but positive moral valence rating, as it could be regarded as a morally good action. No LDT has controlled this additional valence, which might reveal something that was already present in the binocular rivalry experiments: that neutral and positive-morally-valenced words may not differ in RT or Accuracy. With just “morally charged” and “neutral” words categories, this difference cannot be assessed at all. Likewise, the rest of the studies do not seem to be able to control all the factors that stem from a bounded-rationality, decision-making experimental set, to support the symmetry, nor to perform an adequate data analysis for such purposes.

As a final note, it is also important to note that no data analysis can provide a clear cut between correlation and causation. Causality is an abstraction that goes beyond statistical relationships, although being supported by them. It involves developing theoretical frameworks, considering counterfactuals, and making sense of the underlying relational mechanisms between variables. Providing linear models as a standard to establish causality is not sufficient, as the latter demands integrating theoretical consideration—statistical analyses show how certain variables covary with others.

Revisiting AMP: the theoretical side

With the supporting data explained away, let us now move to the conceptual side. In this section, I will address some theoretical caveats that AMP needs to face in order to be a plausible theory. First, I will examine the theoretical prerequisites that AMP entails. Then, I will formulate a different example to avoid the crossfire between the social and moral realm

that lurks behind some of the data AMP relies on. Third, I will try to further develop the AMP argument and try a *reductio ad absurdum*. Finally, I will try to challenge AMP against clinical and developmental data.

Challenging the priors

MP's Richness of perception and cognitive penetrability

It is no doubt that moral properties are high-level properties, and so they are subject to the ongoing debate on whether we represent high-level properties (HLP) in perception. To pave the least possible controversial path towards her theory, the attentional moral perceptualist needs to make a decision on how she addresses this issue.

A first option could be, intuitively, to defend the richness of perception in a broad sense, so moral properties are represented in perception (Siegel, 2006). However, this version of the argument in favour of the richness of perception has to face strong criticism: that the overall perceptual phenomenology of what seems an experience of a HLP is not sensory phenomenology, but the result of a perceptual judgment or some other cognitive episode (Brogaard, 2013); that the overall perceptual phenomenology of what seems an experience of a HLP is due to shifts of attention (Price, 2009); and that the overall perceptual phenomenology of what seems an experience of a HLP is just a gestalt appearance instead of a HLP. Bence Nanay advanced the debate in a plausible way: contra Siegel, he states that the only HLP that are susceptible to be represented in perception are those that are action-oriented (e.g., perceiving an apple as “edible” or a tree as “climbable”). In this case, AMP requires the moral property to be something related to an action. The closest take could be that moral wrongness and rightness prompts us to jump into action, but this is not usually the case—this is, they do not seem to yield strong motivational force. One could also reformulate Nanay’s account and say that the properties we perceive are, instead of morally right/wrong, “*morally reprehensible*” and “*morally praisable*”, but, yet again, none of the related actions tend to happen. Moreover,

in committing to this “translation” of properties, it is difficult to see why can’t *any* property be translated into an action-oriented property, e.g., the beauty of something into “*aesthetically admirable*”. This would be the case in which, for every HLP, there is a possible translation into an action-oriented property. I highly doubt that this is what Nanay referred to as action properties, and, of course, he meant to argue against Siegel’s proposal.

Against the initial intuition expressed above, a third option could be to defend moral properties as low level properties (LLP). This is, without a doubt, a difficult task to do, and, most certainly, one with huge theoretical implications. There are, *prima facie*, two ways in which one could defend this. If, on the one hand, the properties that can be represented in perception (LLP) have been genetically determined by evolution through the development of some cognitive mechanism that is specifically sensitive to them, then the moral perceptualist needs to defend the attunement to moral properties as innate, and at the same level of contrast detection, colour detection, and so on.

Regarding the nativist aspect of this option, it is worth noting that there are some theories, such as the Universal Moral Grammar Theory by Mikhail (2007), and recently revitalized by Carchidi (2020), that have already tried to do the difficult task of explicating the primitives of morality as a hard-wired cognitive mechanism.

On the other hand, if the mechanism that is sensitive to moral properties, again, understood as LLPs, is hard-wired through training instead of being innate, we need to offer an account of how this is possible: what is the nature of this plasticity, and the nature of the training itself?

Although cognitive penetration seems the best asset for such purpose, relying on it will leave the moral perceptualist facing the same objections as CMP, as I mentioned in the Introduction. Compiled transducers (Pylyshyn, 1984), too, might be posed as the explanatory mechanism for this process, for they would allow post-perceptual processing to be “ensembled”

on perception (Pylyshyn, 1999). Nonetheless, it would still be necessary to specify which perceptual mechanisms, and how, are altered by the influence of cognition, questions that seem as difficult to answer as the previous ones.

Overall, it seems that there is no effective way to incorporate cognitive penetration into AMP without, sooner or later, having to face Reiland's (2011) Redundancy Argument. According to him, the philosophical picture of what happens when we witness a morally charged situation would go like this: first, we experience the scene; second, we have an affective empathic response to what we see; third, this lead to us having the moral seeming, intuition or judgment that the act is right or wrong; finally, we endorse the seeming/intuition/judgement and form the belief that the act is right or wrong.

The extra step that AMP proposes, in which the affective empathic response or the seeming/intuition/judgment also cognitively penetrates the visual experience, causing us to see it as right or wrong, does not seem add anything to the story. As Reiland says, "*absent any other reason to postulate it, we shouldn't*" (Reiland, 2021, p. 324).

The moral and the social

Apart from the controversial nature of its priors (richness of perception and cognitive penetrability), there are also problems with AMP's formulation and classic examples, which I will now address.

Most of the examples provided in the MP literature, regardless of the specific variant, are rooted in cases in which the moral rightness/wrongness of the action is certainly aligned with what is socially acceptable/unacceptable to do. Moreover, some paradigmatic cases do not yield any moral weight *per se*, but rather social weight, or even a complex mix of both in which each variable's influence becomes indistinguishable from the other's. For example, while it is true that running over a pedestrian is morally wrong, the driver that does not respect

the red light is not doing something morally wrong *per se*—rather, she is doing something that is socially unacceptable/wrong. The relevance of controlling the ‘social charge’ of a situation in order to control all possible variables may lead us to a debate about the nature of moral properties, in which I advocate for a non-reductionist view of them, essentially orthogonal to social ones (see, e.g., Van Schoelandt, 2018, or Forbes et al., 2020, for a defence of the overlap between the two). In this sense, social properties are related to the “*standards of behaviour that are based on widely shared beliefs about how individual group members ought to behave in a given situation*” (Bernhard et al., 2006, p. 217). The key in the differentiation between the two is precisely where the differentiation of social and moral norms stem from. Schram & Charness (2011, p. 1) say, it is that “*social norms involve observation by others and external sanctions for violations, while moral norms involve introspection and internal sanctions*”. Moreover, another trait of social properties is their cultural specificity—as the name points out, there is no social property without a social group.

The issue is a complex one, but, for my purposes here, I will just stick to these characterizations of social and moral properties, and to orthogonality between the two. This is, I defend that it is plausible to find a situation with either combination: morally and socially wrong; morally and socially right; morally right and socially wrong; morally wrong and socially right³.

For most societies, it is true that our social norms, to a greater or lesser extent, tend to be aligned with humanitarian rights and egalitarian views. Even our legal system is also consonant to what we consider humanitarian rights, so not only “*morally wrong*” behaviours are usually legally punished, but also socially punished. According to the Social Learning

³ While there are moral and social wrongs, it must be said that the property itself of rightness and wrongness is not different—it just stems from different norms in each case (moral norms/social norms).

Theory (SLT), children can learn through the mere observation of another agent action's consequences. It is worth noting that, in humans, social reward and social punishment are one of the greatest incentives to behave one way or another (Faeri & Delgado, 2014). Therefore, it is safe to say that social "rightness" and "wrongness" categories are formed early in life, and result from universal, basic learning mechanisms.

For example, take the humanitarian norm "*not to harm others*". In most societies not only a decontextualised act of violence is morally unacceptable, but also socially punishable. People do not engage with violent people; people who commit violent acts are condemned to an abnormal course of socialization, dominated by estrangement and social rejection/fear. At the same time, it is legally forbidden to execute violence against one another. So, at the end of the day, many systems (moral, socio-cultural ones) are at play in most of the literature that tries to advocate for MP. With the aim to depurate the correct target of AMP, I suggest working with an example in which no other variable, such as legal or social consequences, might yield explanatory power in the alleged perceptual, attentional representation of moral-difference makers.

Example 1. *Richard is a young man living in West Virginia, 1950, when the Jim Crow laws are in force. He is having lunch at a café, but he is suddenly interrupted by a commotion. Richard witnesses how an African American man is kicked out of the place (E1), as it is an "all-white" space. He does not see it as an immoral act (O1), whatsoever. Nevertheless, once he is at home, he decides to reflect on what has happened. He starts thinking about segregation, and, after many headaches, he concludes that it is an unjust situation. The next day, at the same café, Richard witnesses another black man being kicked out of the place (E2). This upsets him, as, by now, he is aware of the unfairness of the situation (O2).*

Undoubtedly, we are now dealing with something that, whilst morally wrong, it is not socially nor legally reprehensible, given the context. Let's keep on with the argument from here.

Now, has anything changed in Richard's perception in the first and second scene in the café?

As I have previously mentioned, a moral perceptualist, regardless of the variant she adheres to, is committed to defend the richness of perception. The argument would go as it follows (adapting Siegel's argument, 2006):

1. *If the overall mental state O1 of which E1 is a part differs in phenomenal character from the overall mental state O2 of which E2 is a part, then there is a phenomenal difference between the sensory experiences E1 and E2.*
2. *If there is a phenomenal difference between the sensory experiences E1 and E2, then E1 and E2 differ in content.*
3. *If there is a difference in content between E1 and E2, it is a difference with respect to the moral properties that are represented in perception.*

In principle, to maintain the richness of perception, the phenomenal difference between overall mental experiences O1 and O2, should not be best explained by other any fact than the perceptual (i.e., the sensory phenomenology) of the property "moral wrongness/rightness". Let us assume that this is the case (premise 1)

However, against premise 2, one could propose that, while there is a phenomenological difference in E1 and E2, it does not arise from a difference in the visually experience, but from a difference in the distribution of perceptual attention. When people develop moral principles, their attention becomes "attuned" to them, hence causing a change in how Es are phenomenologically experienced. This is, precisely, the move that AMP makes to avoid

committing to representation of moral properties —and, consequently, going back to CMP proposals—. In making that move, the argument now would look like this:

I. If the overall mental state $O1^$ of which $E1^*$ is a part differs in phenomenal character from the overall mental state $O2^*$ of which $E2^*$ is a part, then there is a phenomenal difference between the sensory experiences $E1^*$ and $E2^*$.*

II. If there is a phenomenal difference between the sensory experiences $E1^$ and $E2^*$, then $E1^*$ and $E2^*$ differ in the distribution of perceptual attention.*

Still, for the proposal to endorse MP, a third step must be added, that can either be:

III. If there is a difference in the distribution of perceptual attention between $E1^$ and $E2^*$, then it is a difference with respect to the moral salience patterns that are represented in **attention**.*

III. If there is a difference in the distribution of perceptual attention between $E1^*$ and $E2^*$, then it is a difference with respect to the moral salience patterns that are represented in **perception**.*

It cannot be III*, as we would get entangled with the CMP proposal again. If changes in attention prompt changes in the perceptual experience itself (“representation in perception”), then AMT goes all the way back to CMP, which has already faced fair criticism.

So, at this point, the argument would look like this:

I. If the overall mental state $O1^$ of which $E1^*$ is a part differs in phenomenal character from the overall mental state $O2^*$ of which $E2^*$ is a part, then there is a phenomenal difference between the sensory experiences $E1^*$ and $E2^*$.*

II. If there is a phenomenal difference between the sensory experiences $E1^$ and $E2^*$, then $E1^*$ and $E2^*$ differ in the distribution of perceptual attention.*

III. *If there is a difference in the distribution of perceptual attention between $E1^*$ and $E2^*$, then it is a difference with respect to the moral salience patterns that are represented in attention.*

This is quite obscure—what does it mean to represent something at the attentional level, and how does it escape from CMP?

Now, attention itself has a variety of ways to be defined. One of them seems to be particularly fitting for AMP: attention as a way of perceiving, whose phenomenal character is different from that of the attended perceptual experience itself (Mole, 2011).

Intuitively, if we stick to this definition, and add representationalism to the picture, in defending that attention leads to changes in the content of perception, we are committing to AMP as another version of CMP. So, we would need to redefine attention as a state with its own phenomenology, which would mostly consist of rendering a stimulus as salient (Wu, 2011). Still, another question hangs from this definition: if attention has its own phenomenology, and phenomenology is explained in terms of content (as per the representationalism framework is assumed here), what is it that attentional states represent? According to Speaks (2010), there are two possibilities: either they represent properties of the relevant perceptual experience, or they represent aspects of the scene represented by that perceptual experience.

Both, nonetheless, have problematic implications. On the one hand, if they represent aspects of the scene represented by the perceptual experience, then it remains unclear why we need a redundant system to represent what it is already represented in perceptual experience. In the presented example, the attentional state should render the visual moral salience patterns, which consist in, say, the colour of the customer's skin, the 'x-race space' sign (in this case, an all-white space sign), and the aggressiveness of the movements of the man that kicks out the

costumer. But this is already represented in Richard's visual experience! How would attention make any difference?

It seems that the step that AMP must take is towards the second alternative. If attentional states represent properties of the relevant perceptual experience, then, the difference between E1 and E2 is that, in E2, the attentional state represents Richard's visual experience as representing morally salient difference-makers (in both cases Richard's visual experience is representing the same properties, but his attentional state represents his visual experience as representing them as morally salient). In this sense, attentional states would work as higher-order representational states (metarepresentational states) that (meta)represent properties about the first-order representational state (the perceptual experience). However, as Speaks (2010) points out, this is quite odd, as it goes against the general intuition about the transparency of perception: that even when we attend to our perceptual experiences, what we get is what our perceptual experiences represent.

What is more, as Speaks (2010) also argues, even meta-representational states should be able to misrepresent. Yet, attentional states, understood as (meta)representing properties of the perceptual experience fail to do so:

"After all, how could attentional states misrepresent, on the present view? Presumably by representing the relevant perceptual experience as representing something which it does not, in fact, represent. But what would such a state be like" (Speaks, 2010, p. 339).

Suppose that Richard's attentional state (meta)represented his experience of the "all-white-space" sign as representing an all-white-space sign, but the perceptual experience itself did not in fact represent the sign as an "all-white-space" sign, but rather as a "mixed-race-space" sign. Would Richard's total phenomenology (the phenomenology of the experience itself *and* the phenomenology of the attentional state) include the phenomenal character typical of

experiences which represent signs as all-white-spaces signs, or not? If it did, then, contra the hypothesis, Richard's visual experience would be representing the relevant cue, the “all-white-space” sign, as an all-white-space sign. But if it did not, i.e., if the overall phenomenology does not include the phenomenal character typical of experiences which represent signs as “all-white-space” signs, then it would seem that (again, contra the hypothesis) Richard's attentional state is not representing the perceptual experience as representing the sign as an “all-white-space” sign. In other words, attentional states, understood as states that (meta)represent properties of the perceptual experience attended to seem infallible, which leaves the moral perceptualist, undoubtedly, in a highly controversial spot.

Neural substrate and clinical cases

Let's now move on to consider some of the psychological and clinical implications of AMP.

If AMP is true, something akin to an ‘attentional disfunction’ should be at play in those psychopathologies that course with an impoverishment of the performance in moral tasks, for the person would not be able to attend to moral cues and represent morally salient patterns. If this is true, then, one should expect to 1) relate attentional malfunctions to some psychopathologies' symptoms, and 2) show some sort of structural or functional changes within moral development that causally relate to attentional development.

To assess the issue, then, I will hold onto the case of psychopathy, a “*personality disorder involving severe disruption in moral behavior accompanied by pronounced deficits in emotion*” (Glenn et al., 2009, p. 2).

Data with respect to attentional impairments in psychopathic individuals is rather conflictive. While some studies show an actual enhancement of selective attention, others propose a deficit in attentional narrowing, and, others, too, defend just a “*more effortful, top-down attention processing*” (Gao et al., 2009, p. 816). Despite the collision of the hypotheses,

it seems that there are some peculiarities regarding attention in psychopathic population. Nonetheless, attention might not be a causal element in the impoverishment of their performance (i.e., there is not anything ‘wrong’ with attention); instead, the variations in attention might be a *consequence* of other impairments.

In fact, most studies regarding the moral dimension of psychopathic symptomatology target the hypoactivation of the amygdala as a key to the abnormal processing in moral decision tasks, as well as the diminished functional communications between the amygdala and the ventromedial prefrontal cortex to be the real cause of the variability in moral-decision tasks (vmPFC) (Wolf et al., 2015). VmPFC provides information about the personal and subjective value of stimuli or actions, even in decision-making settings where there is no choice (Levy & Glimcher, 2016; Yoder et al., 2016). Its connectivity with other reward regions just stresses its function in value encoding of real and hypothetical objects and behaviours (Xia et al., 2015). As for the amygdala, it has been already addressed how it is a centre of affective evaluation (Zald, 2003). Precisely, neurodevelopmental data reveal that retarded maturation of the amygdala (Gao et al., 2009), as well as early lesions of the vmPFC manifest in severe antisocial behaviour, insensitivity to future consequences of decisions, and repeated failure to respond to behavioural interventions later in development (Bechara et al., 2000).

This suggests an interesting course of analysis: it might not be a deficit in attention, nor ‘moral attention’, but rather in the processing of the value of the attended and perceived cues. Take, for instance, Glass & Newman (2006)’s work on psychopathic offenders’ performance in an emotion recognition task. The task had two sub-sets, one with free attentional exploration, and other with participants’ attention directed to the facial expression. In both conditions, participants’ scores did not significantly differ from the scores of non-psychopathic population, and were able to perform even better than them in the detection of emotion. The tests included fearful expressions, which can also work as a moral difference-maker under specific

circumstances (say, a person A who is given a clown-doll by her close friend B and seems frightened; it is A's expression of fear that prompts us to think that B has acted wrongly). Both in the condition where attention was directed to the facial expression and where it is not, the results held.

The diminishment of attention, then, would happen because of a top-down influence that, as data show, does not impair detection, but rather naturally constricts "*what is worth attending/keep attending*" according to certain pathological expectations. These expectations are pathological in the sense that are built on an abnormal value assignment to certain stimuli, that ultimately lead to a disinterest in them.

Now, let us flip the example and examine if attentional impairments themselves tend to course with, and cause, moral deficits. Several studies on children with Attention-Deficit/Hyperactivity Disorder (ADHD) show that there is not a significant correlation between ADHD and performance on moral-decision tasks (see, e.g., Meier et al., 2007), nor do children with ADHD have an impoverished moral developmental course (Fernandes et al., 2014)

Conclusions

In this essay, I have examined the empirical data that seems to support AMP, trying to discuss both the methodological limitations and the alternative interpretations of the data. By doing this, I have pointed out two fundamental methodological constrains for AMP: the asymmetry of the AMP results and the 'construct validity' of the paradigmatic empirical and theoretical examples. In regard to the latter, I have articulated an example of a moral situation that aims to deplete all the confounding variables, whose analysis has brought up the contentious stances that AMP needs to commit to: the existence of phenomenologically rich 'attentional states', the abandonment of transparency of perception, and the inability to misrepresent. Additionally, I have revised some further implications of AMP, as the reduction of some psychopathological symptoms to attentional deficits.

Hopefully, it has become evident that, while attention is important in processing the stimuli, and it surely relates to perception, the processing of morally charged situations as salient is not likely the result of perceptual attentional mechanisms—instead, the role of attention is more of a bridge between cognition and perception, being cognitive biases, heuristics, expectations, and beliefs the main tenet in this interaction. As Fireston & Scholl (2016, p. 2) said, “*the discovery that people attend to just outcomes in moral scenarios tells us something about what we expect in moral scenarios, but it doesn’t tell us anything new about how attention itself works. We already knew that attention is drawn to what people expect*”.

References

- Anderson, E., Siegel, E. H., & Barrett, L. F. (2011). What you feel influences what you see: The role of affective feelings in resolving binocular rivalry. *Journal of Experimental Social Psychology, 47*(4), 856-860.
- Bandura, A., & Walters, R. H. (1977). *Social learning theory* (Vol. 1). Prentice Hall: Englewood cliffs.
- Bechara, A., Tranel, D., & Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain, 123*(11), 2189-2202.
- Bernhard, H., Fehr, E., & Fischbacher, U. (2006). Group affiliation and altruistic norm enforcement. *American Economic Review, 96*(2), 217-221.
- Carchidi, V. J. (2020). The Nature of Morals: How Universal Moral Grammar Provides the Conceptual Basis for the Universal Declaration of Human Rights. *Human Rights Review, 21*(1), 65-92.
- Chang, H. J., K. Huang, and C. Wu. "Determination of sample size in using central limit theorem for weibull distribution." *International Journal of Information and Management Sciences, Vol. 17, No. 3. 2006, pp. 153-174.*
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159.
- Cowan, R. (2015). Cognitive penetrability and ethical perception. *Review of Philosophy and Psychology, 6*, 665-682.
- Cunningham, W. A., & Brosch, T. (2012). Motivational salience: Amygdala tuning from traits, needs, values, and goals. *Current Directions in Psychological Science, 21*(1), 54-59.
- Fareri, D. S., & Delgado, M. R. (2014). Social rewards and social networks in the human brain. *The Neuroscientist, 20*(4), 387-402.
- Fernandes, A. P. A., Dell'Agli, B. A. V., & Ciasca, S. M. (2014). O sentimento de vergonha em crianças e adolescentes com TDAH. *Psicologia em estudo, 19*, 333-344.
- Firestone, C., & Scholl, B. J. (2015). Enhanced visual awareness for morality and pajamas? Perception vs. memory in 'top-down' effects. *Cognition, 136*, 409-416.

- Firestone, C., & Scholl, B. J. (2016). Moral perception 'reflects neither morality nor perception. *Trends in Cognitive Sciences*, 20(2), 75-76.
- Foot, P. (1967). The problem of abortion and the doctrine of the double effect.
- Forbes, M., Hwang, J. D., Shwartz, V., Sap, M., & Choi, Y. (2020). Social chemistry 101: Learning to reason about social and moral norms. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 653–670, Online, November 2020. Association for Computational Linguistics.
- Gamer, M., Zurowski, B., & Büchel, C. (2010). Different amygdala subregions mediate valence-related and attentional effects of oxytocin in humans. *Proceedings of the National Academy of Sciences*, 107(20), 9400-9405.
- Gao, Y., Glenn, A. L., Schug, R. A., Yang, Y., & Raine, A. (2009). The neurobiology of psychopathy: a neurodevelopmental perspective. *The Canadian Journal of Psychiatry*, 54(12), 813-823.
- Garon, M., Forgeot d'Arc, B., Lavallée, M. M., Estay, E. V., & Beauchamp, M. H. (2018). Visual encoding of social cues contributes to moral reasoning in autism spectrum disorder: an eye-tracking study. *Frontiers in human neuroscience*, 12, 409.
- Gigerenzer, G. (2004). Fast and frugal heuristics: The tools of bounded rationality. *Blackwell handbook of judgment and decision making*, 62, 88.
- Gilligan, C. (1977). In a different voice: Woman's conception of self and morality. *Harvard Educational Review*, 47, 481-517.
- Gilligan, C. (1982). In a different voice: Psychological theory and women's development. Cambridge, MA: Harvard University Press.
- Glass, S. J., & Newman, J. P. (2006). Recognition of facial affect in psychopathic offenders. *Journal of abnormal psychology*, 115(4), 815.

- Glenn, A. L., Raine, A., & Schug, R. A. (2009). The neural correlates of moral decision-making in psychopathy. *Molecular psychiatry*, *14*(1), 5-6.
- Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work?. *Trends in cognitive sciences*, *6*(12), 517-523.
- Hochman, G., Glöckner, A., Fiedler, S., & Ayal, S. (2016). “I can see it in your eyes”: Biased processing and increased arousal in dishonest responses. *Journal of Behavioral Decision Making*, *29*(2-3), 322-335.
- Keidser, G., Naylor, G., Brungart, D. S., Caduff, A., Campos, J., Carlile, S., Carpenter, M. G., Grimm, G., Hohmann, V., Holube, I., Launer, S., Lunner, T., Mehra, R., Rapport, F., Slaney, M., & Smeds, K. (2020). The Quest for Ecological Validity in Hearing Science: What It Is, Why It Matters, and How to Advance It. *Ear and hearing*, *41 Suppl 1*(Suppl 1), 5S–19S.
- Kneer, M., & Hannikainen, I. R. (2020). Triage dilemmas: a window into (ecologically valid) moral cognition.
- Kohlberg, L. (1963). The development of children’s orientation toward a moral order: Sequence in the development of moral thought. *Vita Humana*, *6*, 11–33.
- Levy, D., & Glimcher, P. W. (2016). Common value representation—A neuroeconomics perspective. *Handbook of value: Perspective from economics, neuroscience, philosophy, psychology and sociology*, 85-118.
- McBrayer, J. P. (2010). A limited defense of moral perception. *Philosophical studies*, *149*, 305-320.
- Meier, B. P., Sellbom, M., & Wygant, D. B. (2007). Failing to take the moral high ground: Psychopathy and the vertical representation of morality. *Personality and individual differences*, *43*(4), 757-767.
- Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological bulletin*, *130*(5), 711.

- Mikhail, J. (2007). Universal moral grammar: Theory, evidence and the future. *Trends in cognitive sciences*, 11(4), 143-152.
- Mole, C. (2011). *Attention is cognitive unison: An essay in philosophical psychology*. Oxford University Press.
- Peeters, G. (1971). The positive-negative asymmetry: On cognitive consistency and positivity bias. *European Journal of Social Psychology*, 1(4), 455-474.
- Peters, U. (2022). What is the function of confirmation bias?. *Erkenntnis*, 87(3), 1351-1376.
- Petoft, A., & Abbasi, M. (2022). Children's Criminal Perception; Lessons from Neurolaw. *Child Indicators Research*, 15(5), 1905-1920.
- Phelps, E. A. (2006). Emotion and cognition: insights from studies of the human amygdala. *Annu. Rev. Psychol.*, 57, 27-53.
- Prinzmetal, W., McCool, C., & Park, S. (2005). Attention: reaction time and accuracy reveal different mechanisms. *Journal of Experimental Psychology: General*, 134(1), 73.
- Pylyshyn, Zenon (1999). Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. *Behavioral and brain sciences*, 22(3):341–365.
- Pylyshyn, Zenon W. (1984). *Computation and cognition*. Cambridge University Press.
- Rahal, R. M., Fiedler, S., & De Dreu, C. K. (2020). Prosocial preferences condition decision effort and ingroup biased generosity in intergroup decision-making. *Scientific reports*, 10(1), 1-11.
- Reiland, I. (2021). On experiencing moral properties. *Synthese*, 198, 315-325.
- Rest, J., & Narvaez, D. (2014). The college experience and moral development. In *Handbook of moral behavior and development* (pp. 251-268). Psychology Press.
- Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and social psychology review*, 5(4), 296-320.
- Schram, A., & Charness, G. (2011). Social and moral norms in the laboratory. *UCSB manuscript*.
- Shuman, V., Sander, D., & Scherer, K. R. (2013). Levels of valence. *Frontiers in psychology*, 4, 261.

- Simon, H. A. (1990). Bounded rationality. *Utility and probability*, 15-18.
- Speaks, J. (2010). Attention and intentionalism. *The Philosophical Quarterly*, 60(239): 325–342
- Teoh, Y. Y., Yao, Z., Cunningham, W. A., & Hutcherson, C. A. (2020). Attentional priorities drive effects of time pressure on altruistic choice. *Nature communications*, 11(1), 3534.
- Toribio, J. (2023). *Seeing wrongness*. [Manuscript submitted for publication].
- Unkelbach, C., Alves, H., & Koch, A. (2020). Negativity bias, positivity bias, and valence asymmetries: Explaining the differential processing of positive and negative information. In *Advances in experimental social psychology* (Vol. 62, pp. 115-187). Academic Press.
- Van Bavel, J. J., FeldmanHall, O., & Mende-Siedlecki, P. (2015). The neuroscience of moral cognition: From dual processes to dynamic systems. *Current Opinion in Psychology*, 6, 167-172.
- Van Schoelandt, C. (2018). Moral accountability and social norms. *Social Philosophy and Policy*, 35(1), 217-236.
- Vance, J., & Werner, P. J. (2022). Attentional Moral Perception. *Journal of Moral Philosophy*, 19(5), 501-525.
- Väyrynen, P. (2018). Doubts about moral perception. *Evaluative perception*, 109-128.
- Werner, P. J. (2020). Moral perception. *Philosophy Compass*, 15(1), e12640.
- Wolf, R.C., Pujara, M.S., Motzkin, J.C., Newman, J.P., Kiehl, K.A., Decety, J., Kosson, D.S. and Koenigs, M. (2015), Interpersonal traits of psychopathy linked to reduced integrity of the uncinate fasciculus. *Hum. Brain Mapp.*, 36: 4202-4209.
- Yoder, K. J., Harenski, C., Kiehl, K. A., & Decety, J. (2015). Neural networks underlying implicit and explicit moral evaluations in psychopathy. *Translational psychiatry*, 5(8), e625-e625.