Záró beszámoló

A felelősség elméletei (PD 113198)
Vezető kutató: Réz Anna

Bevezetés

A 2014 szeptember és 2018 december közötti kutatási időszak során – melyből 16 hónapig a projekt szülési szabadság miatt szünetelt – a projekt célkitűzéseinek megfelelően a morális felelősségtulajdonítás olyan aspektusaival foglalkoztam, amelyek a szabadakarat-vita hagyományos kérdésein túlmenő, azoktól részben független problémákat tárnak fel a felelősségtulajdonítás, a hibáztatás és ezen erkölcsi gyakorlatok normatív igazolásával kapcsolatosan. Kutatásaim kiindulópontját a doktori disszertációm központi tézise képezte, mely szerint az erkölcsi felelősségtulajdonítás feltételeinek meghatározása nem egy tisztán metafizikai-cselekvéselméleti kérdésre adott válasz, hanem jellegéből adódóan normatív etikai vállalkozás, amelynek során elsőként arra a kérdésre kell választ adnunk, hogy mi teszi a felelősségtulajdonítás egyes eseteit – legyen szó hibáztatásról, intézményes büntetésről vagy olyasfajta reaktív attitűdök kifejezéséről, amilyen a harag vagy a neheztelés – erkölcsileg igazolttá vagy akár kívánatossá.

Konferenciaelőadásaim és publikációim többségében ezeknek az összefüggéseknek a feltárására vállalkoztam. A kutatási időszak kitűnő alkalmat biztosított számomra arra, hogy tágabb összefüggésekbe tudjam helyezni korábbi kutatási projektjeimet, és megírhassak olyan tanulmányokat, amelyek egy későbbi monográfia gerincéül szolgálhatnak. Egyben hatalmas segítséget nyújtott a doktori hallgatóból fiatal kutatóvá válás időszakában, mert a segítségével ápolni tudtam hazai és nemzetközi szakmai kapcsolataimat.

A továbbiakban a szűken a projekthez kapcsolódó kutatási eredményeimet, illetve mindazon tudományos tevékenységeket fogom rendszerezni, amelyeket a kutatás mellett végeztem. Először (angol és magyar nyelvű) absztraktok formájában bemutatom az általam kutatott témákat, és jelzem, hogy milyen outputokat sikerült ezekhez rendelnem. Az utána következő részben pedig az ezen időszakhoz tartozó konferenciarészvételeket, publikációkat, az oktatói tevékenységet, illetve az egyéb pályázatokban való részvételt sorolom fel. Végül szót ejtek arról is, milyen jövőbeni célokat alapoztam meg a projekt segítségével, és ennek fényében milyen további terveim vannak.

Kutatási témák

A) A kutatáshoz szorosan kapcsolódó témák

The Normative Interpretation: Methodological Consequences (konferenciaelőadás és megírt tanulmány, bírálat alatt)

In *Responsibility and Moral Sentiments* R. J. Wallace argues that determining the conditions of moral responsibility is an essentially normative project: we are searching for those facts which make holding someone morally responsible *fair*[1]. Wallace offers—what he calls—a normative interpretation of what it is to be morally responsible for something. According to Wallace

> (*N*) *S* is morally responsible (for action *x*) if and only if it would be fair to hold *s* morally responsible (for action *x*).[2]

By providing (*N*) Wallace aims to offer a general schema in order to understand the debate between different theories of moral responsibility, rather than a substantive theory of moral responsibility. So as to develop the schema into a theory, we need to define and characterize the concepts involved.

The normative interpretation has been challenged on various grounds. Some argued that the biconditional expressed by (*N*) is false (Smith 2008, Nelkin 2009), others deemed it gratuitous (Vargas 2004) or trivial (Fischer 2008). Rather than arguing for the feasibility of the normative interpretation, in the following, I will discuss some methodological consequences which follow from accepting it.

Metaphysical Uniformity

Most theories of responsibility strive to give a metaphysically unified account—even if they cannot explain all judgments of responsibility exactly alike, they try to show how different conditions can be traced back to a common root or can be regarded as variations or expansions of one and the same idea. This ambition is a natural consequence of the thought that defining the conditions of moral responsibility involves the discovering of a metaphysical property to which judgments of responsibility respond. If, by contrast, we accept that facts about responsibility are fixed by our moral and ethical standards regulating the fairness of responsibility-attribution, we have no reason to suppose that these considerations will add up to a uniform theory. Appreciating the significance of the normative interpretation might lead to theories which are less tidy and elegant than the ones we are accustomed to.

---

[1] A slightly different formulation has been put forward by Manuel Vargas, who labels Strawson's and Wallace's accounts *practice-based*, referring to their contention that "the 'truth maker' for claims about responsibility is some normative feature of responsibility-characteristic practices" (Vargas 2004, p. 225).

[2] Wallace understands the moral appropriateness in question in terms of fairness, so, following Manuel Vargas (2004), I use a slightly modified version of his original definition.

Extension, Fairness, Responsibility-attribution

As I have previously mentioned, the normative interpretation is only a schema which can be filled in in various ways. This schema contains three variables. The first one is the *extension* of responsibility: what are those things for which we can in principle be held responsible? Since we often make moral judgments about others' attitudes and involuntary omissions, we have to consider the possibility that our responsibility extends far over our voluntary actions and omissions.

Second, we have to give an account of the concept of *fairness* in the present context. For instance, while many authors (typically incompatibilists) offer desert and merit-based accounts of fairness (see e.g. Pereboom 2001), others (mainly compatibilists) are more eager to identify unfairness with unreasonableness (Sher 2005).

The third factor to be considered is the nature of the *moral practices* involved. Responsibility-attributing practices are notoriously diverse, ranging from moral appraisal to praise and blame to institutional punishment. As long as we do not have a clear picture of which practices we regard as typical or paradigmatic, and what the practices imply, it is impossible to tell the fairness of *what* should be guaranteed by the theory.

I take it that a theory of responsibility should not only accomplish these tasks in a coherent way, but also has to create a reflective equilibrium between the theory and our ordinary thinking about the extension of responsibility, the concept of fairness and the nature of our moral practices. It might be—and the literature on free will and responsibility makes this claim likely—that this task cannot be completed without being revisionist in at least one of the three aspects.

Fifty Shades of Responsibility

Given what has been told, it seems that applying this schema will lead us to accept that there can be several adequate theories of responsibility which cannot challenge one another. Depending on how we characterize fairness, responsibility-attributing practices and the extension of responsibility, we will accept (sometimes radically) different conditions of responsibility, which will be suitable only relative to the normative notions we have initially chosen. Although there are several methodological and normative restrictions in play which radically limit the possible number of acceptable theories, at the end of the day, the acceptance of the normative interpretation will inevitably lead us to the thought that responsibility has more than one "face".

References

Fischer, J. M. 2012. *Deep Control: Essays on Free Will and Value*, Chapter 8. New York: Oxford University Press.

Nelkin, D. 2009. Responsibility, rational abilities, and two kinds of fairness arguments. *Philosophical Explorations* 12, pp. 151–165.

Pereboom, D. 2001. *Living Without Free Will*. Cambridge: Cambridge University Press.

Sher, G. 2005. Kantian Fairness. *Philosophical Issues* 15, pp. 179–192.

Smith, A. M. 2008. On Being Responsible and Holding Responsible. *The Journal of Ethics* 11, pp. 465–484.

Vargas, M. 2004. Responsibility and the Aims of Theory: Strawson and Revisionism. *Pacific Philosophical Quarterly* 85, pp. 218–241.

Wallace, R. J. 1995. *Responsibility and the Moral Sentiments*. Cambridge: Harvard University Press.

The Puzzle of Involuntary Carelessness (megjelent tanulmány)

It is a general and consensually accepted demand of theories of moral responsibility that they comply with some central and robust intuitions of ours. Common examples go back to Aristotle: we would consider a theory deeply inadequate if it could not explain why we exempt people from responsibility if they are young children or suffer from serious mental illnesses, or if they acted under compulsion or in ignorance. However, once we have accomplished this task, others will arise: most probably our theory will still yield some counterintuitive results once it comes to more complicated cases.

One such group of problematic cases includes involuntary omissions: instances of carelessness, forgetfulness, absent-mindedness, negligence and the like. Examples are numerous: we hold responsible and blameworthy the driver who caused a car accident by not paying sufficient attention; the teenager who forgot to keep her (otherwise sincere) promise to her parents; the babysitter who did not pay heed to one part of a child's dietary restrictions and consequently caused a severe allergic reaction. I assume that the central puzzle about involuntary omissions comes from the acceptance of the following three claims:

(1) We are morally responsible only for those things over which we exercise control.
(2) People do not exercise control when acting carelessly, forgetfully, absent-mindedly, etc.
(3) People are responsible for (at least some of) their involuntary omissions.

It is beyond the scope of this paper to give an analysis of (1) or to argue for its truth—for our present purposes it is enough to accept that due to its intuitive appeal it would be unreasonable to refute it without any further argument.

As we will see, (2) is probably the most often contested claim among the three. Still, it is not hard to see why we tend to think that we do not exercise control over our involuntary omissions. Although there is not any kind of general consensus about the responsibility-relevant notion of control, traditionally the concept of control is tied up with the notions of intentionality, choice and consciousness. These features of paradigmatically responsible conduct are obviously absent in the case of involuntary omissions: not only do we not *choose* or *intend* to forget, not to notice or not to care about certain things—usually we are not even *aware* of our wrongdoing at the time of its happening. Forgetting, not keeping in mind, not noticing and not paying attention essentially involve the lack of awareness of certain facts, considerations or reasons. Whatever we happen to think about the exact conditions of control-execution, involuntary omissions will most probably fail to fulfill those criteria.

In the following I will first present two promising and popular solutions to the puzzle of involuntary omissions, and discuss their virtues and deficiencies. Then I will turn to Ferenc

Huoranszki's treatment of involuntary omissions in *Freedom of the Will: A Conditional Analysis* and point out its advantages over rival accounts. Finally I make an attempt to answer a serious worry concerning the fairness of holding people responsible for things over which they did not exercise actual intentional control.

Implicit Bias and the Limits of Individual Responsibility (megjelent tanulmány)

The paper explores the aims and perplexities of explaining individual moral responsibility for actions influenced by implicit biases toward members of minority groups. First it presents the notorious difficulties of justifying responsibility for such instances of human agency, which apparently lacks either consciousness or direct control or identification with one's value judgments. The author argues that although indirect and attributionist accounts of moral responsibility adequately explain and justify most of the cases where people are held responsible for their implicit biases, in order to explain "alienated" cases (where someone is influenced by implicit biases in spite of their sincere egalitarian commitments) we would have to accept normatively untenable and counterintuitive variantist proposals. The second part of the paper situates the problem in the contemporary political context and raises some general complaints about the political project which aims to erase structural injustice by diminishing individual biases. The author argues that (i) blaming individuals for being influenced by implicit biases is not an effective method of eliminating discriminatory behavior and that (ii) at least in certain contexts 'weeding out' individual biases is insufficient to end structural inequalities. Thus, the paper concludes, there is no (practical) need to extend the scope of individual responsibility for biased behavior to those cases, where this extension requires substantial theoretical sacrifices.

Strawson, Reversal and the Normative Interpretation (konferenciaelőadások)

Peter Frederick Strawson's seminal paper "Freedom and Resentment" has been subject to constant reinterpretations and re-evaluations since its first publication in 1962. One of its major innovations was first pointed out by Gary Watson, who argued that the most original part of Strawson's account concerns the "order of explanation" with regard to being and holding responsible: "Whereas traditional views have taken these attitudes (i.e., reactive attitudes such as guilt or gratitude) to be secondary to seeing others as responsible, to be practical corollaries or emotional side effects of some independently comprehensible belief in responsibility, Strawson's radical claim is that these "reactive attitudes" (…) are constitutive of moral responsibility (Watson 1987, pp. 259)" In a similar vein, R. J. Wallace has argued (1995), elaborating on Strawson's theory, that there are no prior and independent facts about being morally responsible. Facts about responsibility come about by determining the conditions under which it is morally appropriate to hold someone morally responsible (Wallace calls this the normative interpretation of the free will debate).

Recently, Patrick Todd (2016) and David Shoemaker (2017) advanced a different understanding of the Strawsonian approach. While Shoemaker offers a response-dependent account of moral responsibility, and Todd is highly skeptical about that very project, they both agree that the adequate interpretation of the Strawsonian "reversal" should be conceived

analogously to response-dependent accounts of evaluative properties (such as e.g. funny). According to Shoemaker, one is blameworthy, if anger is a *fitting response* to her conduct, where fittingness is determined by standards inherent to the emotion itself.

In my paper I first argue that we should distinguish the methodological-conceptual claim put forward by Wallace (and presumably Watson) from the metaphysical thesis of Shoemaker and Todd. Wallace's account is a general meta-schema to understand the concept of responsibility and the theoretical debates surrounding it. The response-dependent account proposed by Shoemaker, by contrast, is a theory of moral responsibility itself.

Second, by elaborating on the difference between Wallace's concept of the normative interpretation and response-dependent accounts of moral responsibility, we can shed a new light on Strawson's original paper. Although the text itself is notoriously reticent in spelling out the relationship between reactive attitudes, moral demands and moral responsibility, it seems safe to assume that the response-dependent account proposed by Shoemaker is a useful amendment to the Strawsonian approach. Its major advantage is that it helps interpreting and illuminating Strawson's much discussed and debated claim, i.e., that the practice of interpersonal expectations and the attitudes with which we react to other people's compliance or non-compliance with them cannot and should not be "externally" justified. If the normative standards regulating the fittingness of our responsibility-attributing practices (that is, having reactive attitudes) are internal to the practice itself, then it makes no sense to refer to more general metaphysical or normative principles to justify them.

However, for the very same reason it is open to further discussion whether Wallace's normative interpretation would or should be endorsed by a coherent Strawsonian account. Although strictly speaking the normative interpretation does not commit us to a non-naturalistic approach to moral responsibility, its main motivation runs afoul of naturalism. According to Wallace's proposal, at the end of the day our moral norms will determine the conditions under which it is appropriate to hold someone responsible – and we have no reason to suppose that our norms regulating the fairness of responsibility attribution will pick out a naturally or metaphysically uniform set of properties. Thus, Wallace's account leads to a conclusion contrary to Strawson's original discussion: the final justification of our responsibility-attributing practices comes from an "external" source – from general moral norms and principles which regulate the fairness of responsibility-attribution.

References

Shoemaker, D. 2017. Response-Dependent Responsibility; or, A Funny Thing Happened on the Way to Blame. *Philosophical Review*, Vol. 126, pp. 481–527.

Strawson, P. F. 1962. Freedom and Resentment. *Proceedings of the British Academy* 1962, pp. 1–25.

Todd, P. 2016. Strawson, Moral Responsibility, and the "Order of Explanation": An Intervention. *Ethics* 127, pp. 208–240.

Wallace, R. J. 1995. *Responsibility and the Moral Sentiments*. Cambridge: Harvard University Press.

Watson, G. 1987. Responsibility and the Limits of Evil: Variations on a Strawsonian Theme. In: Schoeman, F. (ed.), *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*, pp. 256–286. Cambridge: Cambridge University Press.

Szabadság és felelősségérzet (konferenciaelőadás és publikált tanulmány)

Egyes érzelmek különösen nagy karriert futottak be az erkölcsfilozófia történetében. Jelen tanulmányban arra teszek kísérletet, hogy jellemezzek egy mindezidáig keveset tárgyalt érzelmi és motivációs állapotot, a felelősségérzetet, és körbejárjam, hogy pusztán abból, hogy létezik ez az érzelmi állapot, mi minden következhet normatív etikai elméleteink számára. A válasz a legkevésbé sem kézenfekvő, hiszen az érzelmek számtalan különböző formában és funkcióban jelennek meg a normatív etikában, a metaetikában és az erkölcspszichológiában. Tanulmányom első részében röviden összefoglalom, milyen metaetikai és metodológiai előfeltevések mentén tudjuk beépíteni az érzelmeket etikai elméleteinkbe. Ezután fenomenológiai tulajdonságai és kognitív tartalma alapján jellemzem a felelősségérzet állapotát, és megvizsgálom, mennyiben van összhangban a felelősségérzet tartalma és megélése azokkal az erkölcsi fogalmakkal, amelyeket a szakirodalom már bőségesen számba vett és jellemzett. Amellett fogok érvelni, hogy a felelősségérzet nem azonos sem a bűntudattal, sem a kötelességtudattal, csak távolról kapcsolódik az erkölcsi felelősség bevett fogalmához, és a benne megfogalmazódó kötelességek nem illeszkednek jól sem a deontikus, sem a konzekvencialista kötelességfogalomhoz. Végül – visszatérve az első rész meglátásaihoz – számba veszem, milyen módon hasznosítható a felelősségérzet a normatív etikai elméletek számára.

Necessary Failure – Compatibilist Methods to Handle 'Ought Implies Can' (konferenciaelőadás és megjelenés alatt álló tanulmány)

The 'ought implies can' (OIC) principle should deeply concern every compatibilist author who do not wish to give up (or even wish to elaborate on) deontic ethical principles. And there are many such contemporary theorists: attributionist accounts of responsibility (Scanlon 1998, Smith 2008, Hieronymi 2004), hierarchical theories or real-self views (Watson 1975, Frankfurt 1971), Fischerian semi-compatibilists (Fischer & Ravizza 1998) and all sorts of capacity views (Wallace 1995, Dworkin 2011) are equally sensitive to the claim that no one can be obliged (and consequently to be held responsible afterwards) to do (or refrain to do) what one is unable to do (or refrain to do).[3]

In my presentation I will analyze two dominant compatibilist methods to reconcile compatibilism with 'ought implies can'. The first one, pursued, among others, by Pamela Hieronymi and Angela Smith, is the inflation of blame. Hieronymi and Smith both weaken the link between responsibility and "full-blown" (affectively laden, actively expressed) blame. Rather, they identify the relevant practices of holding responsible by such milder phenomena as unspoken moral evaluation or criticism. By downplaying the moral force and significance of blame (or at least the sort of blame indicative of moral responsibility) they can then persuasively

---

[3] Proponents of the conditional analysis of free will are the obvious and notable exceptions.

argue that it is not unfair to hold people responsible, even if they couldn't do anything whatsoever to avoid doing (or feeling or thinking) as they did.

The second strategy, taken, among others, by R. J. Wallace and Ronald Dworkin, is to explain away the intuitions which support a strong link between responsibility and avoidability. Both Wallace and Dworkin aim to find alternative *moral* principles governing the fairness of holding responsible, which can do an equally good job in explaining and justifying ordinary excuses and exemptions as those supporting OIC. At the end of the day both authors end up with a capacity view, which links responsilibity to general rational and moral abilities instead of the specific ability to perform a given act at a given time.

At the final section I will analyze why both strategies, in spite of their original insights and internal coherence, fail to impress anyone leaning toward incompatibilism. On the one hand, attributionist accounts build their theories on a concept of responsibility which is only marginally relevant to incompatibilist theorists. Wallace and Dworkin, on the other hand, adopt methodologies (the generalization strategy and the interpretive method, respectively) which block major revisions in our morality system, because they implicitly assume that our morality system cannot be fundamentally spoiled.

References

Dworkin, G. 2011. *Justice for Hedgehogs*. Cambridge Mass.: Belknap Press of Harvard University Press.

Fischer, J. M. & Ravizza, M. 1998. *Responsibility and Control: a Theory of Moral Responsibility.* Cambridge: Cambridge University Press.

Frankfurt, H. 1971. Freedom of the Will and the Concept of a Person. *Journal of Philosophy* 68, pp. 5–20.

Hieronymi, P. 2004. The Force and Fairness of Blame. *Philosophical Perspectives* 18, pp. 115–48.

Scanlon, T. 1998. *What We Owe to Each Other*. Cambridge: Harvard University Press.

Smith, A. M. 2008b. On Being Responsible and Holding Responsible. *The Journal of Ethics* 11, pp. 465–484.

Wallace, R. J. 1995. *Responsibility and the Moral Sentiments*. Cambridge: Harvard University Press.

Watson, G. 1975. Free Agency. *Journal of Philosophy* 72, pp. 205—220.

How Should the Fairness of Blame Be Understood?

It is seldom debated in the literature of moral responsibility that responsibility is closely linked to moral blame, but there is substantial disagreement about the exact nature of the connection. The most straightforward way is to claim that being an appropriate target of moral blame is both a necessary and sufficient condition of being moral responsible for something (see e.g. Wallace 1995). However, some authors (see e.g. Smith 2008, Nelkin 2009) have presented convincing counterexamples, where blame or sanction seems intuitively inappropriate despite

the fact that no one would (again, intuitively) dispute the agent's responsibility for the wrongdoing. For example, Angela Smith distinguishes three types of considerations which influence the appropriateness of blame, but not the agent's responsibility: the moral standing of the blamer (it is inappropriate to blame someone for a fault which we ourselves also regularly commit), the significance or seriousness of the fault (it might be inappropriate to blame someone for something morally insignificant or trivial) and the agent's response to her own wrongdoing (it is inappropriate to blame someone, if she already suffers from guilt and self-reproach).

In my presentation I will argue that these objections are based on an inadequate understanding of such fairness claims. In short, my proposal is based on the recognition that when we talk about the fairness of holding responsible, we always occupy the recipient's perspective. The question is not whether an act of blame is fair *tout court*, but whether it is fair *to the agent* to blame her. Blame, ordinarily understood, might impose burdens on the agent for several reasons: it might imply the impairment of the relationship in question; it can morally discredit the blamed person in the eyes of third parties; question her moral integrity or force her to bear the burden of responsibility by feeling guilt, making apology or compensate for the harm done. The central question is whether it is fair to impose such burdens on the agent—and *this* question is arguably independent from the considerations which Smith presents. For instance, it might be unfair of me to blame someone for not having arrived in time, given that I am habitually late. But at the same time it might still be fair to blame her in the sense that she deserves to feel bad about being late. Based on my previous argument I will finally argue that the link between blame and moral responsibility can be captured by the following definition: *S* is morally responsible for action *x* if and only if *S* cannot legitimately reject the burdens involved in blaming her on the basis of *x*.

References

Nelkin, D. 2009. Responsibility, Rational Abilities, and Two Kinds of Fairness Arguments. *Philosophical Explorations* 12, pp. 151–165.

Smith, A. M. 2008. On Being Responsible and Holding Responsible. *The Journal of Ethics* 11, pp. 465–484.

Wallace, R. J. 1995. *Responsibility and the Moral Sentiments*. Cambridge: Harvard University Press.

B) A kutatáshoz nem vagy csak lazán kapcsolódó témák (válogatott)

Philosophers in the Public Discourse (konferenciaelőadás)

To develop the profile of some sort of a public intellectual a philosopher can pursue two main strategies. First, one can take on more or less traditional philosophical issues and present them in an approachable fashion, thus popularizing philosophy as a scientific discipline. Second, one can take a stand on current, directly non-philosophical public debates, while maintaining an essentially philosophical approach to the subject matter. I find this latter method more fruitful and effective in reaching a wider audience and informing them about the merits of philosophical

thinking (here I take it for granted that it is a good thing in itself if at least some philosophers gain wider recognition in the given society). In my talk I will present some conceptual distinctions, methodological assumptions and key messages, which can be usefully brought into the current Hungarian public discourse by a particularly philosophical standpoint.

When Apprehension and Appreciation Come Apart – A Novel Account of Personal Taste (konferenciaelőadás és megírt tanulmány, bírálat alatt, társszerző: Bárány Tibor)

Classical views on taste typically incorporate two different mental activities which take place when being confronted with a work of art: (i) discriminating subtle qualities in the object; i.e., picking out the relevant aesthetic features and apprehending them in the appropriate way (supposedly, this process does not exhaust in apprehending the work's perceivable non-aesthetic properties, but also requires the possession of relational and historical knowledge on the perceiver's part) and (ii) taking pleasure in aesthetically appreciated works of art (affective component). (On classical models of taste see for example Korsmeyer 2005.) Accordingly, in the Humean model the ultimate standard of taste for ordinary people, who have developed a lower degree of delicacy of taste, is the ideal critic, who appreciates every aesthetically relevant properties and her affections are in perfect harmony with her (immaculate and impersonal) aesthetic judgment. (On Hume-influenced views of aesthetic value and appreciation see for example Levinson 2002, Ribeiro 2007, Taliaferro 1990.)

An especially serious challenge of the classical view comes from the supposed motivational state of non-ideal audience. As Jerrold Levinson puts it: „why a person who is not an ideal critic should rationally seek, so far as possible, to exchange the ensemble of artistic objects that currently elicit his or her approval and enjoyment for some other ensemble that is approved and enjoyed by the sort of person he or she is not." (Levinson 2010) In order to explain how the taste of the ideal critic can or should give reason to ordinary people (and thus to have any normative force), most defenders of the standard view embrace aesthetic empiricism as the adequate account of aesthetic value, i.e. the claim that an item's aesthetic value is its power to yield intrinsically valuable experiences (= aesthetic pleasures) when understood correctly (see e.g. Lopes 2015). Aesthetic empiricism, taken together with the assumption that people are (generally or always) motivated to maximize (aesthetic) pleasure, yields the appropriate result: ordinary people should rationally seek and follow the advice of ideal critics.

The concept of an ideal critic has been contested both for epistemic and ethical reasons. As Kieran (2008) pointed out, we are faced with insurmountable epistemic difficulties in conceiving how our ideal counterpart would respond to a given artwork: the ideal critic should be required "to have (temporally and perhaps logically) incompatible sets of dispositional states and capacities". And, as far as ethical worries are concerned, most probably we have some sort of moral obligation to respect individual tastes: "One might even admire the peculiar kind of creativity people sometimes display even if one detests the result" (Eaton 2008).

In this paper we present a different but related problem with the classical view, which calls for the revision of the conception of ideal critics and the corollary thesis of aesthetic empiricism.

Contrary to what seems to be a wide consensus even in the contemporary literature of taste, our everyday experience provides all too many examples where justified aesthetic judgments and affective reactions come apart, often contradicting to one another. It happens frequently that we take pleasure in a work of art which we judge to be aesthetically poor, and it is no less common to deeply apprehend the aesthetic merits of a piece of art, still feeling untouched by it. For instance, if someone is particularly attracted to didactic art, then she will be disposed to appreciate artworks with robust moral content even if they fail to live up to higher aesthetic standards, while remaining emotionally unimpressed by those masterpieces which depict their subject in a morally neutral or distant way. (Note that this emotional reaction does not necessarily reflect any kind of moral complaint or criticism on the viewer's part. One can find the perspective, which the artwork adopts, morally innocuous, while at the same time not taking pleasure in it because of its approach to morality.)

Moreover, such discrepancies between apprehension and appreciation are not only possible (*pace* Hume) – they cannot even be considered as rational or aesthetic failures on the agent's part, which need to be repaired or eliminated (as, for instance, de Sousa suggests when labeling these cases as instances of aesthetic akrasia). Quite on the contrary: according to our suggestion cases when aesthetic judgments and affective reactions do *not* align are constitutive to one's personal taste. These discrepancies can sometimes be explained by mere gut reactions, but more often they seem to be adequate reflections of one's moral and aesthetic outlook and thus they express something important about one's normative identity.

This characterization of personal taste, however, does not fit well with the thesis of aesthetic empiricism, since the later only obscures the fact how personal taste does and should come about. Thus, in the second part of the paper we offer the rough outlines of a conventionalist/quasi-contractualist account of aesthetic value. According to this model, artworks take a (hypothetical) aesthetic offer to their audience by help of several, aesthetic and non-aesthetic means. Accordingly, their aesthetic value is to be determined by two factors: (i) the *quality* of the offer: how actual, universal, complex, novel is it? (ii) the *success* of the offer: does the artwork fulfill its promise?

The conventionalist account of aesthetic value can adequately explain divergences of aesthetic judgment and emotion. It might well be that we judge that a given artwork fully delivers its artistic offer, determined mostly by the standards of its genre or subgenre, but we find the given genre or subgenre dull, outdated or shallow – consequently the artwork leaves us emotionally completely detached. Or, vice versa, we recognize serious artistic failures in fulfilling the artistic offer, but we still greatly appreciate and enjoy the originality or courage revealed in it. Also, the quality of an offer is something we can legitimately and reasonably disagree about, thus creating a broad sphere for irreducibly personal taste, which can play the role it actually has in our life: to express our evaluative commitments and to contribute to our normative identity through the enjoyment of works of art.

References

Eaton 2008: Marcia Muelder Eaton: "Aesthetic Obligations". *The Journal of Aesthetics and ArtCriticism* 66, 1–9.

Kieran 2008: Matthew Kieran: "Why Ideal Critics are not Ideal: Aesthetic Character, Motivation and Value". *British Journal of Aesthetics* 48, 278–294.

Korsmeyer 2005: Carolyn Korsmeyer: "Taste". In Berys Gaut and Dominic McIver Lopes (Eds.), *The Routledge Companion to Aesthetics*, 2nd edition, London and New York: Routledge, 267–277.

Levinson 2002: Jerrold Levinson: "Hume's Standard of Taste: The Real Problem". *Journal of Aesthetics and Art Criticism* 60, 227–238.

Levinson 2010: Jerrold Levinson: „Artistic Worth and Personal Taste". *Journal of Aesthetics and Art Criticism* 68, 225–233.

Lopes 2015: Dominic McIver Lopes: „Aesthetic Experts, Guides to Value". *Journal of Aesthetics and Art Criticism* 73, 235–246

Ribeiro 2007: Brian Ribeiro: "Hume's Standard of Taste and the *De Gustibus* Sceptic". *British Journal of aesthetics* 47, 16–28.

Taliaferro 1990: Charles Taliaferro: "The Ideal Aesthetic Observer Revisited". *British Journal of Aesthetics* 30, 1–13.

Aristotelian Ethics Without Virtues

Neo-Aristotelian ethical theories revitalized Aristotelian thought by focusing on the role of virtues in ethical conduct. Contemporary virtue ethical theories, however, have been contested on two bases. First, situationist challenges and other findings in social psychology questioned the role of character traits in explaining moral behaviour and the efficacy of rational reflection in forming better character (Doris 2002). Second, virtue ethical theories are often criticised for failing to give proper action-guidance (Van Zyl 2009) and, as a related matter, for not being well-applicable for policy-making.

In my presentation I would like to argue that we can develop an ethical theory which is faithful to the spirit of the Aristotelian theory, yet bypasses the situationist challenge and convey straightforward answers for applied ethical and political questions.

The heart of Aristotle's ethic is the idea that living a happy, self-fulfilling life (*eudaimonia*) is the ultimate goal of human beings, and that happiness, in turn, consists in living in accordance with the function (*ergon*) of human beings. He identifies this function with virtuous activity, that is, activity caused by the rational soul in accordance with virtue.

It is this central notion of human function and its relation to human flourishing which I take to be the fundamental starting point. Even without the concept of virtue we can make a good sense of the interrelated concepts of human function and happiness as having a dual status: on the one hand, they present a normative ideal humans can strive for, but on the other hand they are based on the actual biological, psychic and social functioning of human beings. Such an interpretation would serve not so much as a theory of individual ethical conduct, but it would rather determine

the ingredients and boundaries of good life and thus convey a general picture about how human societies should function.

To take a closer look at human functioning and the conditions of well-being and to use these empirical insights as the basis of our normative ideals can give fairly clear-cut ethical answers to some notorious applied ethical problems. For instance, for some such theory it is a significant fact of human life that human life begins and comes to an end, and that in early childhood and in our last elderly years we are all in need of intensive caring. That this is an unavoidable human condition provides ethical reason for all of us to become competent carers (on a personal level) and also to build such social systems (on a political level), where care work is visible, respected and available for everyone. Similarly, in an Aristotelian framework we can directly argue for the inherent wrongness of pornography and prostitution, by appealing to a normative ideal of human intimacy and sexuality, which these practices betray.

References

Doris, J. M. (2002) *Lack of Character: Personality and Moral Behavior*. Cambridge University Press, Cambridge.

Van Zyl, L. (2009) Agent-based Virtue Ethics and the Problem of Action Guidance. *Journal of Moral Philosophy* 6(1).

Konferenciaelőadások

2015. február 15.
Philosophy of Philosophy Conference, Eötvös József Collegium
Philosophers in the Public Discourse (meghívott előadó)

2015. április 13.
MTA Filozófiai Intézet
A felelősség elmélete (meghívott előadó)

2015. május 11.
Responsibility – The Epistemic Dimension II workshop, VU Amsterdam
Comments on Matt King

2015. május 21.
Cselekvés, autonómia és felelősség konferencia, Pázmány Péter Katolikus Egyetem
Morális felelősség és a normatív interpretáció (plenáris előadó)

2015. szeptember 30.
IX Interuniversity Workshop on Art, Mind and Morality: Art and Negative Emotions
Universidad de Murcia
You Shouldn't Like This. Emotions and Personal Taste (társszerző: Bárány Tibor)

2015. november 20.
Doktoranduszok Országos Szövetsége
Az abortusz mint filozófiai probléma (kommentár)

2016. február 19.
Erasmus Kollégium
Van-e ízlése az ideális kritikusnak? (meghívott előadó)

2016. február 25.
Érzelmek konferencia
MTA Filozófiai Intézet
Ideális kritikusok vs. érzelmes befogadók (előadó)

2016. november 25.
Moholy-Nagy Művészeti Egyetem, Elméleti Intézet
Mi a baj a nemi sztereotípiákkal? (meghívott előadó)

2017. május 15.
MTA Filozófiai Intézet
Mire kell nekünk a kvóta? (meghívott előadó)

2017. május 25.
European Society of Aesthetics Conference/ESA
Freie Universität Berlin
When Apprehension and Appreciation Come Apart – A Novel Account of Personal Taste

2017. június 8.
Ethical Theory and Moral Practice 20th Anniversary Conference/ETMP
University of Pavia
The Normative Interpretation: Methodological Consequences

2017. december 20.
University of Linkoping
Strawson, Reversal and Concepts of Responsibility (meghívott előadó)

2018. március 1.
A fenomenológiai tényektől a filozófiai elméletekig
MTA Filozófiai Intézet
Szabadság és felelősségérzet (meghívott előadó)

2018. március 26.
BME Filozófia és Tudománytörténet Tanszék
Miért érzünk felelősséget? (meghívott előadó)

2018. május 28.

TEDx Budapest, CEU
Korlátozott racionalitás = korlátozott moralitás?

2018. június 13
5th Annual Conference of the European Philosophical Society for the Study of
Emotions/EPSSE
University of Tallin
Reactive sentiments and fitting attitudes

2018. október 18.
Virtue Ethics and Psychology Conference
Róma
Aristotelian Ethics Without Virtues

2018. október 26.
Workshop on Free Will and Epistemology
MTA Filozófiai Intézet
A Compatibilist (and Kantian) Account of Obligations

2018. november 24.
Magyar Filozófiai Társaság
Alkalmazott etika és anyaggyűjtés (meghívott előadó)

2018. december 7.
A pejoratív kifejezések szemantikája és pragmatikája, BME Szociológia és Kommunikáció
Tanszék
Mi számít szexistának?

Publikációk

Compatibilism and 'Ought Implies Can' (megjelenés alatt)

European Journal of Analytic Philosophy, 2019

Felelősség és felelősségérzet

In: Bernáth, László; Márton, Miklós (szerk.) *Tudomány vagy tapasztalat? A fenomenológia és a naturalizmus szerepe a filozófiai elméletalkotásban*, p. 169. Fakultás Könyvkiadó, Budapest, 2018.

Mindannyian szexisták vagyunk

Műút Online: https://www.muut.hu/archivum/25502 (2017)

Szabad akarat újratöltve

In: Bárány, Tibor (szerk.) Kiskáté : *Kortárs filozófiai kiskönyvtár*, p. 214. Miskolc, Szépmesterségek Alapítvány, 2017

Mindannyian szexisták vagyunk

In: Bárány, Tibor (szerk.) *Kiskáté: Kortárs filozófiai kiskönyvtár*, p. 93. Miskolc, Szépmesterségek Alapítvány, 2017

Implicit előítéletek és az egyéni felelősség korlátai

Magyar Filozófiai Szemle 2017/1, pp. 50-67.

Nőiesebb formára hozni a világot – Virginia Woolf feminizmusa

Hévíz 2 (2016), p. 179.

Szabad akarat újratöltve

Műút Online: http://www.muut.hu/archivum/13931 (2015)

Nem elég feminista?

Színház Folyóirat 2014/1, p.

The Puzzle of Involuntary Carelessness

Magyar Filozófiai Szemle 2014/1, p. 34.

Oktatás

2015 tavasz
Erkölcsi felelősség és meta-etika
Eötvös József Collegium

2015 ősz
Introduction to the Philosophy of Social Sciences
ELTE BTK, Filozófiai Intézet

2017 tavasz
A morális felelősség elmélete és gyakorlata
Széchenyi István Szakkollégium

2018 tavasz
Hogy vannak a dolgok, és hogyan kellene, hogy legyenek?
Széchenyi István Szakkollégium

2018 tavasz
Applied Ethics
Moholy-Nagy Művészeti Egyetem

2018 ősz
Alkalmazott etika előadás
ELTE BTK Filozófiai Intézet

2018 ősz
Alkalmazott etika szeminárium
ELTE BTK Filozófiai Intézet

Egyéb pályázati részvétel és intézményes státus

2015 március -: egyetemi adjunktus, Általános Filozófia Tanszék, Filozófiai Intézet, ELTE BTK

NKFI 120375: *Önértelmezés, érzelmek és narrativitás* – társ kutató

Felsőoktatási Intézményi Kiválósági Program (2018-2021, 3 év): *Önvezető autók, automatizáció, normativitás: Logikai és etikai kutatások* – külső kutató

További tervek

A 2015-ben elnyert adjunktusi pályázatomnak köszönhetően a jövőben fennakadás nélkül tudom folytatni megkezdett kutatásaimat. Mivel a nemzetközi peer-reviewed folyóiratokban való publikálásban nem tudtam komolyabb áttörést elérni, folytatni fogom eddig megírt tanulmányaim tökéletesítését, és publikációra küldését, középtávon pedig felhasználni őket egy, a morális felelősség témájában megírt monográfia megírásához. E mellett eddigi eredményeimet a következő két évben az *Önvezető autók, automatizáció, normativitás: Logikai és etikai kutatások* projektben fogom felhasználni, amelynek a következő évtől már hivatalosan is társ kutatója leszek. Ezen kívül meghívást kaptam a Lund Gothenburg Responsibility Projecttől, hogy a következő két akadémiai év egyikében félévig meghívott kutatóként és óraadóként a Lundi Egyetemen kutassak. Amennyiben ezt családi körülményeim lehetővé teszik, és sikerül megfelelő finanszírozási szisztémát kidolgoznunk, mindenképp szeretnék élni ezzel a nagyszerű lehetőséggel.