UNIVERSIDADE DE LISBOA

FACULDADE DE PSICOLOGIA

SCIENCE CAN EXPLAIN OTHER PEOPLE'S MINDS, BUT NOT MINE: SELF-OTHER DIFFERENCES IN BELIEFS ABOUT SCIENCE

Rogério Paulo Fernandes Gouveia

MESTRADO INTEGRADO EM PSICOLOGIA

Área de Especialização em Cognição Social Aplicada

Ano Letivo

2018/2019

UNIVERSIDADE DE LISBOA

FACULDADE DE PSICOLOGIA

SCIENCE CAN EXPLAIN OTHER PEOPLE'S MINDS, BUT NOT MINE: SELF-OTHER DIFFERENCES IN BELIEFS ABOUT SCIENCE

Rogério Paulo Fernandes Gouveia

Dissertação orientada pelo Professor Doutor André Otelo Paraíba Mata

MESTRADO INTEGRADO EM PSICOLOGIA

Área de Especialização em Cognição Social Aplicada

Ano Letivo

2018/2019

Disclaimer

A shorter version of this thesis was submitted to a scientific journal.

Author Note

Somewhere during my academic experience, I stumbled upon a quote from Alexander (1990; see also, Humphrey, 1976). The quote, as cited by Flinn, Geary and Ward (2005), was:

> "... humans had in some unique fashion become so ecologically dominant that they in effect became their own principal hostile force of nature, explicitly in regard to evolutionary changes in the human psyche and social behavior...the real challenge in the human environment throughout history that affected the evolution of the intellect was not climate, weather, food shortages, or parasites—not even predators. Rather, it was the necessity of dealing continually with our fellow humans in social circumstances that became ever more complex and unpredictable as the human line evolved. Social cleverness, especially through success in competition achieved by cooperation, becomes paramount...nothing would select more potently for increased social intelligence...than a within-species co-evolutionary arms race in which success depended on effectiveness in social competition."

At the time I was very much into evolutionary theory and tried to fit it into every piece of research I laid my hands on. This quote seemed to encompass many of the things I was learning throughout the classes about humans, what makes us so different from other species, and why we do what we do. Truth be told, evolutionary theory seems to be able to pierce through just about anything (Dennett, 1995, p. 521; cited by Sommers & Rosenberg, 2003). Alexander (1990) and his comments, in the context of a class on the above-average effect and social comparison, presented a story from a distant past on why that effect came to be around and so consistently found in research of social psychologists. An ever more complex social competition laid out the conditions for effects like better-than-average to

appear, as this relentless battle with others and one's own reality seems to call for, at least, the belief of being something more than one might actually be.

In *The Rebel*, Albert Camus (1956) wrote: "Man is the only creature who refuses to be what he is." It seems as if man is either cognitively wired or highly motivated to believe he is better than what he really is (i.e., to self-enhance) and to believe he is better than others. About this general fictitious reality that people seem to live in, Becker (1971; cited by Aarssen, 2018) stated:

"The world of human aspiration is largely fictitious, and if we do not understand this, we understand nothing about man. …If you reveal the fictional nature of culture you deprive life of its heroic meaning because the only way one can function as a hero is within the symbolic fiction. If you strip away the fiction man is reduced to his basic physical existence—he becomes an animal like any other animal."

The fictitious nature of beliefs allows people to deal with the holes one might (forcefully) get to peek through from time to time. These holes reveal the unknown and the unknowable. But they also reveal the dark reality one is trying hard not to face. Both linger behind the curtains. Very much like blood coming out of an open wound must be properly taken care of, so too one's own fictitious nature cares protection from those who tell us that our thoughts or beliefs do not fit what reality is. As if telling us this wound is not properly sealed. As if more blood was coming out, blood we do not wish to see. And, indeed, against the beliefs and rationalisms of others (like science), which question the veracity and validity of our own sealing materials and our ability to seal the wound, one must find a way to protect himself. For that, people simply perceive others' scope of knowledge as limited, even if it is not the case, or dismiss what could be relevant information, when perceived as opposing theirs.

> «We tend to resolve our perplexity arising out of the experience that other people see the world differently than we see it ourselves by declaring that those others, in consequence of some basic intellectual and moral defect, are unable to see the things "as they really are" and to react to them "in a normal way." We thus imply, of course, that things are in fact as we see them and that our ways are the normal ways» (Ichheiser, 1949, p. 39; cited by Pronin, Gilovich & Ross, 2004).

Note that with this collection of thoughts I am not trying to argue for the value of evolutionary theory. Nor is my objective to claim to have found the reason people share these above-average beliefs. I am simply trying to convey the point that there are many ways one might come about the conclusions that we so frequently get as the results of our scientific experiments. Many were the cases where the poetic or philosophical intuitions of some, non-intently inspired others to investigate through the scientific methods those very possibilities. And there were many times where those very intuitions, as reduced to essence as they were, turned out not to be proven wrong. The same as happened the other way around. Even though sometimes those findings became intuitions that were not quite what science meant for them to be perceived as. Nonetheless, they were knowledge, and people were learning them. And what is the point of science if not, with time, to become fully intuitive for scientists, then no longer scientists, and for people?

In the very first draft of this thesis, I wrote:

"What seems to be a fairly intuitive thought comes to mind (…) From intuition vs. deliberation, hot cognition vs. cold cognition, mind vs. matter, irrational vs. rational, purpose vs. mechanism, even to motive vs. cognition and, still within Psychology, to id vs. super-ego or unconscious vs. conscious, it seems that this impulsive, somewhat spontaneous child-like entity naively navigating itself fearlessly filled with desire and curiosity goes hand to hand with this rule-based, thoughtful and serious looking character, very much like an experienced

father who seems to be highly aware of the other as it holds the ship just up to that one moment the first one, now older and wiser, is ready to go by its own, right back to Eden. That one moment that never comes."

With this, very cryptically, I was trying to draw the world behind and within the dilemma of what is exactly is the role of science, if not to become the father, leading us back to Eden. That very same place where every answer to any problem simply roles off the tongue, no more thinking, as all the thinking has been done, every problem has been solved. Eden is where father and son get together as one, where both sides of a duality are no longer opposites but one thing. Consciousness was the very thing that threw us out of Eden, ashamed, aware of our nakedness we cannot simply go back, the world is now divided in two – the good and the bad, the good in the good and the good in the bad, the bad in the good and the bad in the bad. As the world got divided in halves, so those halves got divided again, and it is our job to keep looking for where exactly is the middle of that one half we then proceed to slice, again and again. Maybe one day we will forget, we have been slicing forever. Or everything is so small, it will feel like whole again. This is, to me, the world science is embedded in. It will come a time where our sons have grown old, and we have grown old as well. Then, we must trust to have taught everything that our sons needed to them themselves become the father. Then, one must not to not let go of the seat. Destiny always comes, and like Oedipus, the son will always come for his father. To not let go of the sit, is to die in pain, to live in Hell. I hope for Science, a place in Heaven.

Resumo

As pessoas têm teorias leigas sobre a mente e sobre como esta funciona (e.g., Dweck, 1999), e também têm teorias sobre o que a ciência é capaz de explicar em relação a essa mesma mente (Fernandez-Duque, 2017; Gottlieb, 2018; Gottlieb & Lombrozo, 2017).

Gottlieb e Lombrozo (2017) procuraram perceber quais eram as intuições que as pessoas tinham em relação ao potencial explicativo da ciência a respeito de vários fenómenos (e.g., amor romântico, moralidade, utilização de perceção e motricidade para alcançar objetos). Para isso, perguntaram sobre os vários fenómenos quanto é que as pessoas achavam que a ciência era capaz de explicá-los totalmente e quanto é que uma explicação científica geraria desconforto. Juntaram a isto uma exploração de potenciais moderadores que pudessem caracterizar os fenómenos como mais ou menos explicáveis e mais ou menos fora do domínio da ciência. Entre estes fatores estavam: acesso através da perspetiva de primeira-pessoa, ou seja, que só a pessoa pudesse conhecer de facto a própria experiência; introspeção, ou seja, a pessoa poder conhecer a experiência através da consulta dos seus próprios pensamentos e das suas próprias emoções; tornar os humanos únicos e excecionais, ou seja, que ser capaz de uma certa faculdade mental é o que torna os humanos excecionais e que só eles têm essa habilidade; envolver controlo consciente, ou seja, que as pessoas têm vontade consciente sobre certo fenómeno e podem influenciar deliberadamente quando, como e porque é que certo fenómeno acontece. Os resultados indicaram que as explicações científicas são impossíveis e desconfortáveis quando um, ou vários, destes fatores caracterizam os fenómenos em causa. Isto até aconteceu para fenómenos que as pessoas não viveram e sobre os quais reconheciam que a sua imaginação não seria capaz de saber como era viver esses fenómenos (e.g., ser surdo e usar pela primeira vez um aparelho auditivo; Gottlieb, 2018), rejeitando a possibilidade de a ciência poder contribuir para compreender e decidir sobre viver ou não a experiência em causa, quando comparado com a própria

imaginação. Mais ainda, quando consideraram os testemunhos de outras pessoas, que tinham vivido os fenómenos em causa, como adequados para saber sobre as novas experiências, preferiram usar a sua própria imaginação na tomada de decisão sobre avançar ou não para o experimentar viver essas experiências. Como sugeriu a autora, pode ser o caso que as pessoas são resistentes à possibilidade de as experiências dos outros ou de as explicações científicas serem aplicáveis a si (Gottlieb, 2018), da mesma forma e na mesma medida em que poderão ser aplicáveis aos outros.

De facto, as pessoas pensam nos outros como menos complexos (Schroeder, Waytz & Epley, 2017; Waytz, Schroeder & Epley, 2014) e parecem acreditar nisto em praticamente todos os domínios (e.g., Alicke & Govorun, 2005; Dunning, Heath & Suls, 2004; Williams & Gilovich, 2008). Condutores envolvidos em acidentes que os hospitalizaram acharam serem tão bons como condutores sem acidentes (Preston & Harris, 1965) e até presos acham serem melhores que outros presos e não-presos em vários aspetos do *self* (Sedikides, Meek, Alicke & Taylor 2014). Ou seja, homens que cometeram diversos crimes e condutores com registo de acidentes graves, acharam ser melhores que outros, mostrando ignorar as suas circunstâncias pessoais e a realidade que é estarem presos ou terem sido hospitalizados.

Para além destes, outros estudos mostraram que as pessoas acreditam, por exemplo, que o seu passado e o seu futuro são mais imprevisíveis que o passado e o futuro dos outros, que o futuro delas tem mais caminhos possíveis e que o seu comportamento é, de forma única, guiado por intenções e desejos (Pronin & Kugler, 2010); ou que os outros são mais enviesados (Kruger & Gilovich, 1999; Pronin, Lin & Ross, 2002).

De particular importância, as pessoas acreditam que os estados internos dos outros são mais acessíveis que os seus (Pronin, Kruger, Savtisky & Ross, 2001). Quando estão a tentar fazer sentido de si mesmas e dos seus comportamentos, as pessoas valorizam muito a introspeção (Pronin, 2009), mas acham que a introspeção dos outros é um fator de

enviesamento. Mesmo que só uma ilusão (*introspection illusion*; ver Nisbett & Wilson, 1977), a consideração pela própria introspeção é muito alta, especialmente para compreender o próprio e o complexo *self* que caracteriza a sua identidade.

Posto isto, a previsão da investigação aqui apresentada é que as pessoas rejeitam a ideia que a ciência consiga explicar como é que elas pensam e o que é que sentem, quando os fenómenos em questão estão associados com um elevado potencial introspetivo. Porém, as pessoas devem estar mais dispostas a aceitar que a ciência explique melhor esses mesmo fenómenos quando estes se passam na mente dos outros, em comparação com dizerem respeito ao próprio.

Os primeiros dois estudos forneceram testes correlacionais de vários potenciais moderadores desta diferença eu-outro: introspeção (de agora em diante, o fator a ter em conta), ser único, controlo consciente, e envolvimento (Gottlieb & Lombrozo, 2017). Os resultados foram os esperados: a diferença eu-outro emergiu moderada pelo fator introspeção. O terceiro estudo oferece evidência experimental direta sobre o efeito desse fator na diferença eu-outro.

Assim, os resultados sugerem que as pessoas têm teorias diferentes sobre o que a ciência consegue explicar sobre si *versus* aquilo que a ciência consegue explicar sobre os outros. Para faculdades mentais fortemente associadas a introspeção, as pessoas acreditam que a ciência explica melhor os outros do que a si próprias.

Algumas limitações são discutidas. Sobretudo, o Estudo 3 carece de algum rigor metodológico e são dadas algumas sugestões sobre como contornar alguns dos problemas, bem como complementar o trabalho realizado com outros estudos potencialmente relevantes para compreender mais profundamente o fenómeno.

Na discussão, procurou-se integrar os resultados com algumas hipóteses clássicas em Psicologia Social. Primeiro, a hipótese do ator-observador sugere que, como ator, o foco da

atenção está no conhecimento privado que o próprio tem sobre os seus desejos e as suas intenções em relação a aspetos situacionais aos quais a pessoa procura responder, enquanto que, como observador, o foco está no comportamento observável do ator, não tendo acesso a pistas privadas (Pronin, Gilovich & Ross, 2004). No entanto, esta hipótese não esclarece todas as dúvidas que existem quanto aos nossos resultados, bem como os dos outros trabalhos sobre diferenças eu-outro. Existe um longo debate na literatura destas diferenças entre a perspetiva motivacional e perspetiva não-motivacional (e.g. Chambers & Windschitl, 2004; Sherman & Sherman, 1999). A primeira sugere que os vieses comparativos estão ao serviço de objetivos relevantes para o próprio e outros motivos. A segunda sugere que os vieses comparativos são a consequência de características das tarefas e dos mecanismos cognitivos. Nesse sentido, são apresentadas brevemente as duas posições e discute-se a possibilidade de interação entre as duas formas de vieses.

O trabalho termina com uma discussão das implicações práticas dos resultados, bem como da importância destes resultados para o papel da ciência e da comunicação científica.

Palavras-chave: epistemologia leiga; teorias intuitivas; diferenças eu-outro; introspeção

Abstract

Three studies show that people differ in their lay beliefs concerning the degree to which science can explain how their mind works versus how the mind of other people works. In particular, people are more receptive to the idea that the psychology of other people is explainable by science than to the possibility of science explaining their own psychology. This self-other difference is moderated by the degree to which people associate a certain mental phenomenon with introspection.

Keywords: folk epistemology; intuitive theories; self-other differences; introspection

Table of Contents

Introduction 1

Lay Theories 1

The case of science. 2

Self-other Differences 3

The case of introspection. 5

Our Research 7

Experiment 1 8

Method 8

Participants 8

Design, procedure and measures. 8

Results 10

The direct effect of target. 10

Moderation by introspection. 11

Moderation by uniqueness/special. 12

Moderation by control. 13

Moderation by involvement. 14

Discussion 14

Experiment 2 15

Method 15

Participants 15

Design, procedure and measures. 15

Results 17

Moderation by introspection. 17

Moderation by control and feeling special. 18

Discussion 18

Experiment 3 18

Method 18

Participants 18

Design, procedure and measures. 19

Results 20

General Discussion 20

References 27

Science Can Explain Other People's Minds, But Not Mine: Self-Other Differences in Beliefs about Science

Introduction

As a student of psychology, I understand now to have committed one of the most terrible sins a student of such crafts can commit: I took psychology into my household. As scientific evidences drooled out of my mouth to simply explain why my family was doing whatever it was doing, I was received with, what appeared to me at the time, a very defensive tone and an open rejection of the knowledge I well-intently was giving. That knowledge might very well serve others, but not them, they said. Only, when that knowledge proved to be congruent with their theories, a conversation about those very evidences was possible. I am now aware of the terrible mistake I made but became very intrigued as to what exactly went on upon their rejection.

Lay Theories

Before those events, I understood that the knowledge people have about the world is a compilation of thoughts that help them make sense of complex and ambiguous behaviors they face daily. These informal explanations can be called lay theories. Amongst these are theories about the mind and how it works (e.g., Dweck, 1999). There is evidence that people use their knowledge of situational pressures to make predictions about the behaviors of others (though not their own), showing some form of theory as to how it plays out in various contextual settings (Balcetis & Dunning, 2013). It has also been shown that those who believe that personal attributes are malleable (incremental theorists) and those who believe in fixed traits (entity theorists) differ in the way they make predictions about behavior, as well as the way they make trait inferences (Chiu, Hong & Dweck, 1997). Furthermore, people have lay theories of happiness (Furnham & Cheng, 2000), emotion (Labroo & Mukhopadhyay, 2009;

Ong, Zaki & Goodman, 2015; Tamir, John, Srivastava & Gross, 2007; Wilson & Gilbert, 2003), suicide (Knight, Furnham & Lester, 2000), and the causes of alcoholism (Furnham & Lowick, 1984), among many others (see Furnham, 1988).

The case of science.

In addition to those many theories that people have about the mind and how it works, people also have certain intuitions about what science can and cannot explain about the mind (Fernandez-Duque, 2017; Gottlieb & Lombrozo, 2017; Lévy-Leboyer, 1988; Weisberg, Keil, Goodstein, Rawson & Gray, 2008).

Regarding this, Gottlieb & Lombrozo (2017) wondered about the intuitive beliefs that people have concerning the limits of science in explaining the human mind. They asked, for several phenomenon (e.g., romantic love, morality, using perception and motricity to grab an object), how much people believed that science could fully explain them and how much a scientific explanation would generate discomfort. They also identified factors that characterize phenomena as "less explainable" and out of the scope of science. Among these were: first-person access, meaning that only the person can know exactly her own experience; introspection, meaning that the person can know her own experience through the examination of her own internal thoughts and feelings, making humans unique/exceptional, meaning that being able to do a certain phenomenon is what makes humans exceptional and that only they have that ability; conscious control, meaning that people have conscious will over said phenomenon and can deliberatively influence when, how, or why it happens. Their results were as expected: Scientific explanations are impossible and uncomfortable when one, some, or all those factors characterize the phenomena being judged.

But the limits of science go beyond lived personal experience (Gottlieb, 2018). In fact, when people are deciding whether or not to go under new transformative (and non-

transformative) experiences, there is evidence that they consider their imagination to be the most appropriate source of decision, over the testimony of others and scientific evidence. This happens despite people claiming that imagination was the worst way to know what new experiences would be like. To them, science was regarded as bad at knowing what new experiences would be like and equally poor as a source for decision on whether or not to engage in said experiences. These considerations were related to how much an experience was perceived as personally transformative. So, even when people know that they could not know what the consequences of undergoing such experiences will be, they prefer to rely on their own imagination than trusting science or others.

In fact, even when the testimony of others was considered to be the best at knowing what a new experience would be like, people's own imagination was still favored over the experiences of others when it came to the decision-making process. As said by the author, a possibility could be that people are resistant to the idea of others' experiences or of scientific generalizations being applicable to them (for a more extended debate see Gottlieb, 2018), the same way those same generalizations might be applicable to others.

Self-other Differences

People tend to see others as less complex (Schroeder, Waytz & Epley, 2017; Waytz, Schroeder & Epley, 2014) than themselves, possibly making people more accepting of the possibility that science might explain others' minds better than their own complex minds.

There is a long and extensive literature in social psychology on the perceived differences between the self and others. It is generally shown that people see themselves as better than others in just about anything (e.g., Alicke & Govorun, 2005; Dunning, Heath & Suls, 2004; Williams & Gilovich, 2008). The thing is, not everyone can be better than most. That defeats the whole concept of average. Yet, most people believe to be better-than-

average, putting themselves in a better light than others (intently or non-intently; the debate about motivated and non-motivated accounts of biases can be found in the General Discussion).

Along the tradition of research on self-other differences it has been found that, for instance, people believe that others are more susceptible to persuasive tactics (Davison, 1983), more motivated by extrinsic vs. intrinsic incentives (Heath, 1999), more conformist in domains ranging from consumer purchases to political views (Pronin, Berger & Molouki, 2007), and less free (Helzer & Dunning, 2012; Pronin & Kugler, 2010).

Pronin and Kugler (2010) showed that people believe their pasts and futures to be less predictable than others' pasts and futures, that their future had more possible paths and that their own future behavior was uniquely driven by intentions and desires. And this last result has been somewhat shown before: people seem to believe that their intentions are important to make assessments about their own behavior but fail to find it equally important to make assessments about others (Pronin, 2008).

People also think that other people are more biased than they are (Kruger & Gilovich, 1999; Pronin, Lin & Ross, 2002). In Kruger and Gilovich's (1999) study, whether participants were married couples, video game enthusiasts, debaters or dart players, they expected others to be biased when allocating responsibilities for a series of desirable and undesirable joint outcomes. And the same is true when comparing themselves to the average American, classmates in a seminar, and other airport travelers (Pronin et al., 2002). In this last study, their less biased status persisted even after reading a description on how they could have been affected by said bias, showing to dismiss information opposing their beliefs or to reject other's opinions and theories about the workings of their minds.

Another example of this could be, for example, people believing to be better drivers (Preston & Harris, 1965; Svenson, 1981). Preston and Harris (1965) compared drivers with

accident reports, whose driving involved them in accidents serious enough to require hospitalization, with drivers without accident reports. They were matched in sex, age, race, education level and not distinguishable on performances on written tests. Only the number of previous traffic violations separated them. The ones with accidents, when describing themselves on a driving performance continuum, rated themselves closer to the "expert" than the "very poor" level, similarly to the "safe" drivers.

Even prisoners, when questioned about their moral and ethical characteristics, believe to be better than the average prisoner and even the average non-prisoner (Sedikides, Meek, Alicke & Taylor, 2014). Prisoners were asked to judge themselves on how, for example, moral, trustworthy, honest, generous, self-controlled and law-abiding they were compared to others. Even when comparing themselves to non-prisoners, they believed to be better than the average person, only believing otherwise with respect to a single trait: law-abidingness. This means that men who committed various crimes, independently of the severity of the crime, widely shared the belief of being better than the average person, dismissing their own personal circumstances and the palpable reality of being in prison or hospitalized by an accident.

The case of introspection.

As reviewed in the previous section, there are many ways in which others are seen as limited compared to us (e.g., Alicke, 1985; Alicke, Klotz, Breitenbecher, Yurak & Vredenburg, 1995; Allison, Messick & Goethals, 1989; Dunning, Meyerowitz & Holzberg, 1989; Messick, Bloom, Boldizar & Samuelson, 1985; for a review, e.g., Alicke & Govorun, 2005).

Part of those limitations might be what makes the testimonies of others to be less relevant than our own imagination during the decision-making process previously presented

(Gottlieb, 2018), or what makes people see science as incapable of explaining them on certain phenomena (Gottlieb, 2018; Gottlieb & Lombrozo, 2017). We know that people think that others' minds are more biased and constrained, as well as less complex than theirs. And we also know that people believe that others' inner states and dispositions are knowable whereas theirs are less accessible (Pronin, Kruger, Savtisky & Ross, 2001). This accessibility problem can also be true for science, thus leading people to think that the minds of others are more easily explained by science, whereas theirs are not so easily reducible to general principles that apply to everyone.

Of most importance for the present proposal is the finding that people value introspection highly when trying to make sense of themselves, their judgments and behaviors (Pronin, 2009), whereas the same does not hold when thinking about other's introspection (on the contrary, they think that introspection is a biasing factor).

So, we hold highly positive beliefs about our introspection's ability to assess why we did what we did and whether we were contaminated by bias or not, but (again) not when judging others' introspections (Pronin et al., 2004). This asymmetry persists even after people are given access to the introspection of others (Pronin & Kugler, 2007), once again supporting their own personal "unbiased" beliefs over what is then dismissed (and possibly relevant) information.

This asymmetry is referred to as the introspection illusion. The case is that, although people are able to give accurate information about the contents that might come about through introspection, the same cannot be said about their ability to accurately comprehend the processes generating those thoughts and guiding their behavior, since they are mostly inaccessible (Nisbett & Wilson, 1977; a finding discussed and revised by e.g., Johansson, Hall, Sikström & Olsson, 2005; Petitmengin, Remillieux, Cahour & Carter-Thomas, 2013; Smith & Miller, 1978).

Nevertheless, people still hold their own introspection as very capable and as of great importance if they want to get closer to understanding what they really are and the intricacies of their complex self. The more a phenomenon is considered to be introspective the lesser it is seen as explainable by science (Gottlieb & Lombrozo, 2017).

Our Research

Given the research on what people believe to be the limits of scientific explanations and the very prevalent self-other differences, our goal with this study was simple: to explore self-other differences in people's notions about the extent to which science can explain psychological phenomena.

The prediction is that science is not seen as capable of explaining psychological phenomena to which people have exclusive introspective access and which people therefore think that they and only they can understand. Therefore, introspection should be a critical moderator in people's intuitions concerning what science can explain about their minds (Gottlieb & Lombrozo, 2017; Gottlieb, 2018) and the minds of others. People believe that only they can understand what they think and how they feel, by looking inward and consulting their thoughts and feelings, as they are less accessible (Pronin et al., 2001).

So, our hypothesis is that people reject the idea that science can explain how they think and what they feel, when the psychological phenomena in question are associated with a strong feeling of introspection. However, we suspect that people are more willing to accept that science can explain those very same phenomena when they take place in the mind of others. So, for highly introspective phenomena, people are likely to believe that the minds of others will be better explained by science than theirs.

To test this hypothesis, Studies 1-2 test whether people believe that science can explain certain aspects of mental life in themselves versus other people. Moreover, those

studies provide correlational tests of several potential moderators of this self-other difference: introspection, uniqueness, control, and involvement. To anticipate the results, the predicted self-other difference emerged, and introspection moderated this difference. Having identified introspection as a moderator in the first studies, Study 3 offers direct experimental evidence for its effect on shaping beliefs about the reach of science for self-versus-others.

Experiment 1

Method

For all three studies, all measures, manipulations, and exclusions are disclosed, as well as the method of determining the final sample size, and the data were not analyzed before collection was completed.

This research project was approved by the ethics board of the Scientific Council of the Psychology Department of the University of Lisbon.

Participants.

The sample size was not pre-determined. Rather, participants responded voluntarily to a request sent out to a mailing list from a large European university. The number of valid responses was 276 (64% female, *M* age = 28.67, *SD* = 12.75).

Design, procedure and measures.

Participants were informed that the study was concerned with what people think of what science can and cannot explain about the way the mind works, and how people think and feel. Participants were randomly assigned to one of two conditions: In the self-first condition (n = 141), they were told that "*This study is interested in what you think about what science can explain about your mind. You will read about different domains of life, and we want you to indicate to what extent you believe that science can explain the way you think, feel and behave for each of them"*, and then they answered the first block of dependent

variables. The second block was similar to the first one, except that now participants were told: "*At this stage, this study is interested in what you think about what science can explain about the mind of other people in general. You will read about different domains of life, and we want you to indicate to what extent you believe that science can explain the way other people think, feel and behave for each of them*" (the dependent variables in this second block were similar to those in the first block). For the others-first condition (n = 135), the order of these instructions was reversed.

In both blocks, participants rated (from 1 – not at all, to 9 – totally) their level of belief in, and discomfort with, the possibility that science can explain each of seven different domains (i.e., two items – belief and discomfort – for seven domains, multiplied over two target blocks).

Because each domain appeared twice (as there were two blocks), the way that it was presented differed slightly across blocks, so as to avoid a consistency bias in the way people answered for self vs. others. For instance, if in one block participants expressed their belief that science could explain how people fall in love, in the other block they were asked about whether can explain how people feel romantic feelings. The other six domains were: personal relations and social psychology (why people become friends with certain people / why people belong to certain groups); likes (why people like certain music / why people like certain books); judgment and decision-making (why people make certain decisions / why people make certain moral judgments); interpersonal feelings (why people feel empathy/compassion)*;* politics (why people have certain political ideologies / why people vote for certain parties); spirituality and religion (why people are religious or not / how people experience spirituality). Although these seven domains were presented within-subjects twice, they were always presented in a random order within each block.

Afterwards, and for each domain, participants were asked to express their agreement (from 1 – not at all, to 9 – totally) with the following statements, designed to assess potential moderators: 1) there were two items assessing introspection (e.g., *only a person him or herself can understand how he/she falls in love and experiences romantic feelings; a person who is in love and experiencing romantic feelings can know what he/she is feeling through introspection*); 2) two items assessing uniqueness and feeling special: (e.g., *if science were able to explain the way in which a person falls in love and experiences romantic feelings, that would make the love and feelings of that person less special; … that would make that person less unique and more similar to others*); 3) two items assessing control (e.g., *if science were able to explain the way in which a person falls in love and experiences romantic feelings, this would mean that these things are out of his/her control; …that these things are pre-determined and there is no free-will*); and 4) two items assessing involvement (e.g., *How important is for you to be in love and experience romantic feelings?; To what degree are being in love and experiencing romantic feelings fundamental things in your life?*).

Results

Analyses used linear mixed models controlling for the repeated effects of block (block 1 vs. block 2) and domain. For random effects, we included intercepts for subject and domain to adjust for possible variation.

The direct effect of target.

Belief in science. Target (self vs. others in general), block (block 1 vs. block 2), target order (self-other vs. other-self) and block order (block 1-2 vs. block 2-1) were entered as fixed effects. Results showed a main effect of target (*estimate* = -0.15, *SE* = 0.04, $t(1790.83) = -3.50$, $p < .001$, 95% CI [-0.24; -0.07]). The effects of block, target order and block order were not significant (all *t*s < -1.15; all *p*s > .252). Simple pairwise comparisons revealed that

belief in science was lower for the self (M = 5.27, SE = 0.11) than for other people in general (M = 5.43, SE = 0.11).

Discomfort with science. No direct effects of the independent variables emerged as significant for the second dependent variable: discomfort with science (all ts < -1.59, all ps > .111).

Moderation by introspection.

Belief in science. We entered target (self vs. others in general), block (block 1 vs. block 2), target order (self-other vs. other-self) and block order (block 1-2 vs. block 2-1) as fixed effects. We entered the dimension of introspection as the moderator, followed by the interaction term of target by introspection. The results yielded a main effect of target (*estimate* = -0.15, SE = 0.04, t(1786.71) = -3.50, p < .001, 95% CI [-0.24; -0.07]), qualified by a moderation effect of target by introspection (*estimate* = -0.10, SE = 0.04, t(1786.70) = -2.19, p = .029, 95% CI [-0.18; -0.01]). No other main effects were significant (all ts < -1.23, all ps > .217).

Decomposing this interaction effect, results show that when participants believed that a certain phenomenon was highly associated with introspection, the explanatory ability of science was thought to be lower for the self (M = 5.21, SE = 0.16) than for others (M = 5.46, SE = 0.16, F(1, 1786.71) = 16.30, p < .001); when that was not the case, the perceived explanatory ability of science did not differ for self (M = 5.33, SE = 0.16) and others (M = 5.39, SE = 0.16, F < 1, p = .349).

Discomfort with science. We repeated the previous analysis using discomfort as the dependent variable. The effect of introspection was significant (*estimate* = 0.77, SE = 0.11, t(289.69) = 7.14, p < .001, 95% CI [0.56; 0.98]), qualified by an interaction effect of target by introspection (*estimate* = -0.12, SE = 0.04, t(1738.24) = -3.19, p = .001, 95% CI [-0.20; -0.05]; the other effects were not significant, ts < -1.58, ps > .114).

The interaction effect shows that when participants believed that a phenomenon was associated with introspection and could only be understood by oneself, they reported a lower level of discomfort with the notion that science can explain that phenomenon for the self (M = 3.42, SE = 0.15) as compared to others in general (M = 3.60, SE = 0.15, $F(1, 1738.26)$ = 11.41, p = .001). When a phenomenon was not highly associated with introspection, the reported discomfort was equivalent for self (M = 2.12, SE = 0.15) and others (M = 2.06, SE = 0.15; $F(1, 1738.24)$ = 1.26, p = .262).

Moderation by uniqueness/special.

We repeated the same analysis using as a moderator the extent to which participants felt that if science could explain their psychology, this would make them less unique and special.

Belief in science. The results showed a main effect of target (*estimate* = -0.15, SE = 0.04, $t(1786.62)$ = -3.49, p < .001, 95% CI [-0.24; -0.07]), a main effect of uniqueness (*estimate* = -0.41, SE = 0.11, $t(295.25)$ = -3.79, p < .001, 95% CI [-0.62; -0.20]), both qualified by a marginal moderation effect of target by uniqueness (*estimate* = -0.08, SE = 0.04, $t(1786.60)$ = -1.87, p = .061, 95% CI [-0.17; 0.004]). No other main effects were significant (all ts < -1.31, all ps > .193).

When decomposing this interaction effect, results show that when participants believed that scientifically explaining a phenomenon would make a person less unique and special, the perceived explanatory value of science was lower for the self (M = 4.79, SE = 0.15) than for others in general (M = 5.02, SE = 0.15, $F(1, 1786.62)$ = 14.53, p < .001). When participants did not believe that explaining a phenomenon would undermine one's uniqueness or feeling special, there were no differences in belief in science for self (M = 5.77, SE = 0.15) versus others (M = 5.84, SE = 0.15, $F(1, 1786.60)$ = 1.33, p = .249).

Discomfort with science. When discomfort was the dependent variable, there was a main effect of uniqueness (*estimate* = 1.02, *SE* = 0.10, *t*(292.20) = 10.08, *p* < .001, 95% CI [0.82; 1.22]), qualified by an interaction effect of target by uniqueness (*estimate* = -0.12, *SE* = 0.04, *t*(1740.46) = -3.05, *p* = .002, 95% CI [-0.19; -0.04]). The remaining effects were not significant, all *t*s < -1.57, all *p*s > .116.

The interaction effect shows that when participants believed that if science could explain a certain phenomenon about someone that make that someone less unique and special, they reported less discomfort with the explanatory power of science for the self (*M* = 3.67, *SE* = 0.14) when compared to others in general (*M* = 3.85, *SE* = 0.14, *F*(1, 1740.48) = 10.75, *p* = .001). When participants did not think that a scientific explanation would compromise uniqueness or feeling special, discomfort with science did not differ for self (*M* = 1.86, *SE* = 0.14) versus others (*M* = 1.80, *SE* = 0.14; *F*(1, 1740.46) = 1.06, *p* = .304).

Moderation by control.

We repeated the same analyses using control as moderator.

Belief in science. There was a main effect of target (*estimate* = -0.15, *SE* = 0.04, *t*(1787.81) = -3.49, *p* < .001, 95% CI [-0.24; -0.07]), such that science was perceived as less capable of explaining the self than other people in general. No other main effects or interaction effect were significant (all *t*s < -1.38, all *p*s > .168).

Discomfort with science. There was a main effect of control (*estimate* = 0.91, *SE* = 0.10, *t*(290.77) = 8.67, *p* < .001, 95% CI [0.70; 1.12]), qualified by an interaction of target by control (*estimate* = -0.14, *SE* = 0.04, *t*(1736.94) = -3.64, *p* < .001, 95% CI [-0.22; -0.07]; the remaining effects were not significant, all *t*s < -1.57, all *p*s > .117).

Breaking down the interaction, when participants believed that if science could explain a phenomenon this would imply less control and free will, they reported less discomfort with the possibility of a scientific explanation about the self (*M* = 3.54, *SE* = 0.15)

than about others (M = 3.74, SE = 0.15, F(1, 1736.96) = 13.66, p < .001). When participants did not think that a scientific explanation would imply less control, they did not differ in their discomfort with the idea of science explaining themselves (M = 2.00, SE = 0.15) or others (M = 1.92, SE = 0.15; F(1, 1736.94) = 2.10, p = .147).

Moderation by involvement.

Finally, we repeated these analyses testing involvement as a moderator.

Belief in science. The only significant main effect was the effect of target (*estimate* = -0.15, SE = 0.04, t(1789.45) = -3.49, p < .001, 95% CI [-0.24; -0.07]), such that science was again seen as better able to explain other people than oneself. All other effects were not significant, all ts < -1.14, all ps > .253.

Discomfort with science. There was a main effect of involvement (*estimate* = 0.31, SE = 0.04, t(1733.54) = 2.75, p = .006, 95% CI [0.09; 0.54]), such that the idea of science explaining a certain psychological phenomenon was more uncomfortable the more involved people were with that phenomena, but no other effects were significant (all ts < -1.59, all ps > .113).

Discussion

The results of Study 1 were at best mixed and far from conclusive. On the one hand, the results with the first dependent variable (belief in science) are as predicted: in general, people find it easier to believe that science can explain the psychology of other people than their own, and this in part seems to relate to some of the moderators that were explored, such as associating a certain phenomenon with introspection, or thinking that scientific explanations make a person feel less unique and special. However, the results with the other dependent variable (discomfort with science) sometimes went in the opposite direction. A posteriori, one can think of reasons for this incongruity, such as the possibility that one is not uncomfortable with the idea that science can explain something about oneself precisely

because one does not believe that it can do so. Still, this explanation is tentative at best, and these results call for further investigation. The following studies seek to provide better tests of the predictions, using much fewer items (Study 1 was quite long and possibly tiresome and confusing in its many measures) and simpler instructions.

Experiment 2

Method

Participants.

We used G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007) to conduct a power analysis based on fixed effects (main and interaction effects) for *F*-tests. Accordingly, and in order to capture a medium-to-large effect ($f = 0.30$), with $\alpha = .05$, power = .80, for two different groups, the required sample size was 90 participants. To have a sample robust to attrition, we recruited a few additional participants: 100 participants, from which we excluded four who reported having participated in Experiment 1. Thus, our final sample size was 96 (*M* age = 35.86, *SD* = 13.34; 79% female), recruited in exchange for course credit.

Design, procedure and measures.

Target (self vs. other) was manipulated between-subjects. Participants were randomly allocated to one of the two target conditions (self, $n = 49$; other, $n = 47$), where they were specifically given instructions to imagine a scenario for the self vs. someone else. Participants read the following:

> "This study is about the beliefs that people hold about how science can explain people's behavior, thoughts, and feelings. Imagine that a team of scientists made an exhaustive study about how ***you/someone else*** think. This is the most exhaustive study ever done about how someone's mind works. The study involves an exhaustive amount of tests, including techniques of brain imaging (mapping the activated areas

of the brain when ***you/someone else*** think(s)), physiological measures (heart rate, skin conductance, etc.), clinical interviews and psychometric tests.

At the end of the study, these scientists are confident that they can fully understand the way how ***you/someone else*** think(s). Specifically, they claim to be able to explain how ***you/someone else*** feel(s) and think(s) about several life domains: for example, the romantic choices that ***you/the other person*** make(s) and how ***you/the other person*** fall(s) in love and feel(s) love; why ***you/the other person*** have(has) the friends you/(s)he have(has) and belong(s) to certain groups; why ***you/the other person*** like(s) certain music and certain books; how ***you/another person*** feel(s) empathy and compassion for others in suffering; and why ***you/the other person*** are(is) or are(is) not religious and how you/(s)he experience(s) spirituality. Additionally, these scientists claim that, based on these results, it is possible to predict what ***you/someone else*** are(is) going to do, think and feel in the future. For example, based on these results, it is possible to predict the type of person with whom ***you/the other person*** are(is) going to fall in love."

Afterwards participants completed three items assessing their belief in the explanatory power of science, and then three other items assessing potential moderators (presented in random order).

Dependent variable.

The perceived explanatory power of science was measured using three items: "*To what extent do you believe the conclusions of this study?*"; "*To what extent do you believe that this is possible, that is, that a scientific study is able to fully understand how* ***you/someone else*** *think(s)?*"; "*Even if you believe that science is still not able to fully understand how* ***you/someone else*** *think(s) currently, do you believe that this will be possible*

one day?". Responses were given on a scale from 1 (not at all) to 9 (totally). Aggregating the 3 items, $\alpha = .87$.

Potential moderators.

Then participants read: "*Regardless of whether you believe in this or not, we now ask you to think about what it would mean if this was possible, that is, admitting that it was possible. If a scientific study were able to fully understand how* ***you/another person*** *think(s), what would be the implications of such a study?*" Next, they rated their agreement with each of the following statements, from 1 (totally disagree) to 9 (totally agree).

Control. "*If this were possible, it would mean that* ***you are/another person is*** *not truly in control of what you/(s)he think(s) and do(es), that is, that* ***you/another*** *person are(is) not truly responsible for your(her/his) thoughts and feelings and that these things are pre-determined*".

Special. "*If this were possible, it would mean that there is nothing special about how* ***you/another person*** *think(s) and feel(s), and that all your(her/his) thoughts and feelings are merely the result of biological mechanisms operating in the brain*".

Introspection. "*Only* ***you/another person*** *can truly understand how you/(s)he think(s) and feel(s); someone else on the outside can try to understand, but (s)he will never be able to truly know how you/(s)he think(s) and feel(s)*".

Results

Moderation by introspection.

We conducted a univariate ANOVA, entering explanatory power of science as the dependent variable, target as the factor, introspection as the moderator, and the respective interaction term. Results yielded a significant interaction effect of target by introspection, $F(1, 96) = 5.54$, $p = .021$, $\eta_p^2 = .06$. The main effects of target and introspection were not significant (Fs < 1, ps > .496, η_p^2s < .005).

Decomposing the interaction effect, simple pairwise comparisons show that when introspection was rated more highly (+ 1 *SD*) participants believed that science had less explanatory power for the self (*M* = 5.14, *SE* = 0.37) than for another person (*M* = 6.22, *SE* = 0.34, *F*(1, 92) = 4.70, *p* = .033, η_p^2 = .05). When introspection was rated lower (- 1 *SD*) participants did not differ in their beliefs about science for self (*M* = 6.15, *SE* = 0.32) vs. others (*M* = 5.55, *SE* = 0.39, *F*(1, 92) = 1.40, *p* = .240, η_p^2 = .02).

Moderation by control and feeling special.

We repeated the same analyses for each of the other two potential moderators, but the resulting effects were not significant (*F*s < 1, *p*s > .356, ηp^2s < .009).

Discussion

Summing up what we found so far, the only consistent result across the two previous studies is the moderating effect of introspection in the self-other difference in beliefs about science. Study 3 narrows down on introspection and provides a more direct, experimental test of its effect on thinking about science for oneself versus others.

Experiment 3

Method

Participants.

We used again G*Power 3.1 (Faul et al., 2007) to conduct a power analysis to determine the sample size needed to detect an effect using only within-subjects' factors. Based on the interaction effect of interest in Experiment 2, with *f* = 0.22, α = .05, and power = .80, the required sample size was 43. We recruited 45 participants (*M* age = 24.82, *SD* = 5.13; 51% female) in the Prolific Academic platform.

Design, procedure and measures.

Introspection was manipulated within-subjects. Specifically, participants were asked to consider ten different mental phenomena, which were selected from Gottlieb and Lombrozo (2017) as scoring high or low in phenomenological introspection (we adapted items from among the highest- or lowest-ranking introspection scores averaging across Gottlieb & Lombrozo's Studies 2a and 3a-b; in these studies, items were rated on a scale from 1, strongly disagree, to 7, strongly agree and the average across studies will be presented next following each item). The high-introspection phenomena were: 1) Feeling compassion for those in suffering, $M = 5.67$; 2) The love one feels for one's parents, siblings or children (adapted), $M = 5.76$; 3) The ability to appreciate a beautiful landscape, $M = 5.5$; 4) Feelings of spirituality $M = 5.49$; and 5) Falling in love, $M = 5.64$. The low-introspection phenomena were: 1) Being better or worse at recognizing people (adapted), $M = 3.56$; 2) Making certain mathematical errors, $M = 4.36$; 3) The ability to perceive depth and calculate distance (adapted), $M = 4.37$; 4) Being better or worse at reading a map (adapted), $M = 4.53$; and 5) How people combine perception and motricity to reach objects (adapted), $M = 4.72$. High- and low-introspective phenomena were rated above and below, respectively, the midpoint of the scale.

Participants were given the following instructions: "Imagine that a team of scientists did a study to try to understand how people e.g., *feel compassion for those in suffering*." Next, participants were asked to indicate how much they agreed with three sentences (from 1 = totally disagree, to 9 = totally agree):1) "*Although this study would be able to explain how other people in general e.g.,* ***feel compassion for those in suffering****, it would not be able to explain how it happens with me*"; 2) "*The results of this study would help to understand how others e.g.,* ***feel compassion for those in suffering****, but they would never really be able to get to the bottom of how I feel and think about it*"; and 3) "*Even though I believe that scientific*

*studies can make sense of how some people e.g., **feel compassion for those in suffering**, I doubt that they could explain how I truly think and feel about it*". To be consistent with Experiments 1-2, we reverse-scored the three items so that higher values mean higher perceived explanatory power of science for the self.

Each of the ten events was randomly presented to participants, and each of them was followed by these three items.

Results

We conducted a Linear Mixed Model analysis, controlling for the repeated effects of event and introspection. Additionally, for random effects we included the intercept for subject to adjust for possible variation. Results revealed a main effect of introspection on the explanatory power of science, $b = 1.30$, $SE = 0.14$, $t(365.04) = 9.38$, $p < .001$, 95% CI [1.03; 1.57]. When considering high-introspection events participants judged that science had a lower explanatory power for self vs. others ($M = 4.98$, $SE = 0.20$) than when considering low-introspection events ($M = 6.28$, $SE = 0.19$).

General Discussion

These studies suggest that people have different beliefs regarding what science can explain about the way they think versus the way other people think. Study 1 showed that, in general, people see science as better able to explain the psychology of other people than their own, and that this is particularly the case when a certain psychological phenomenon is highly associated with introspection (though there were other significant moderators in this study, and results were not consistent across dependent variables).

Study 2 replicated this interaction, whereby science is seen as having a greater explanatory power for other people than for oneself, and that this is particularly the case when introspection is involved. Whereas Studies 1-2 provided correlational evidence, Study 3

provided an experimental test of the role of introspection in self-other differences in thinking about science and what it can explain. The results gave clear support to those of the previous studies: For highly introspective phenomena people believe that science is better at making sense of others than of themselves, whereas this self-other difference disappears when introspection is not thought to be involved.

There are, of course, limitations to these studies. In Study 3, for example, one could argue that the design is quite simple and that the meaning of participants' answers is not readily clear. To illustrate, when answering "*Although this study would be able to explain how other people in general* [low-introspective phenomenon]*, it would not be able to explain how it happens with me*", disagreeing could mean either that one believes that science explains oneself better than others, or that science explains both equally well (or poorly). But we have no reason to believe the former, as our other two studies as well as those of others (e.g., Pronin, Kruger, Savtisky & Ross, 2001; Schroeder, Waytz, & Epley, 2017; Waytz, Schroeder & Epley, 2014) suggest that that is not a likely hypothesis.

Moreover, and still concerning Study 3, we understand that the qualitative meaning of the average answer for highly introspective phenomena to mean neither agreeing nor disagreeing. Thus, people on average are not saying that others are better explained than themselves. It is our belief that the first two studies can offer some arguments in this regard, enabling us to disentangle the possible meaning of this lack of position-taking. Why would people not disagree (for highly introspective phenomena) with the very same sentences that they had no trouble disagreeing with when asked for low introspective phenomena? It is our suggestion that there is something very different about those phenomena, which makes oneself less explainable that others. Science seems to be limited not so much by the complexity of phenomena but rather by the very first-person perspective that one must have to understand high introspective phenomena fully. This is a great challenge for science, given

that scientific methodology implies in its very concept a third-person perspective. So, people are not exactly agreeing with the items, but they do not disagree because, of all the three possible ways their answer could go (*agree*, *disagree*, or neither), others being better explained than them is something they could not disagree with. Future studies might help to further clarify this point by presenting a forced-choice paradigm, where people either agree or disagree, thus avoiding the neither-kind of answer, and possibly accounting for how much time they take to answer or how confident they feel about their choice.

The reason why people seemingly believe to be less explainable when it comes to highly introspective phenomena deserves further investigation as well. It could be the case that the effects found in the present study are a consequence of people perceiving themselves as more complex than others (Schroeder et al., 2017; Waytz et al., 2014). Furthermore, the results of Pronin and colleagues (2001) on the illusion of asymmetric insight suggest that people might be considering observable behavior (e.g., interpersonal revelations) as a better indicator of what others might be, contrasting with private thoughts and feelings as better indicators of what themselves might be. This would add that piece of "unexplainable" possibly guiding the self-other difference effect that we have been describing. People seem to believe that those very thoughts and feelings are part of what best describes their true nature (Markus, 1983) and these types of bias, as aforementioned, are consistently found throughout research on self-other differences (e.g., Alicke & Govorun, 2005; Dunning et al., 2004; Williams & Gilovich, 2008).

Relatedly, it is perhaps relevant to discuss here the actor–observer hypothesis (Jones & Nisbett, 1972). This hypothesis states that "people tend to explain their own behavior with situation causes and other people's behavior with person causes" (Malle, 2006, p. 895). The fact that situational causes might be what better explains our behavior certainly makes a case for the lesser predictability quality of our past and future (Pronin and Kugler, 2010).

Situations being so different from one another might be what makes our private thoughts and feelings so mysterious and inaccessible (Pronin et al., 2001), as our reactions to very specific and unpredictable situations might be as specific and unpredictable as those. Still, this relation is not quite clear considering that this asymmetry is found for negative events, but the opposite is true for positive events (Malle, 2006).

What could be the case is that, as an actor, one's focus of attention is on one's private knowledge of desires and intentions in relation to the situational features that one must account for. Whilst the observer, having no access to those private clues, can mostly focus on the observable behavior of the actor (as explained in Pronin, Gilovich & Ross, 2004). So, when judging how explainable one is compared to others, it could be that, non-intently, one does not account for everything else other than the observable behaviors of others when making assumptions about their explicability. The focus of attention is simply shifting depending on actor-observer perspective.

Despite this being a candidate account for what is happening, it does not however solve the problem as to why it is happening. Supposedly, people learn about others being as much mental beings as they are from a very early age (see e.g., Astington, 1993). The "theory of mind" hypothesis posits that we understand others as having their own mental states, thoughts, wants, motives and feelings (even chimpanzees seem to have some understanding of this; Premack & Woodruff, 1978). So, it is not quite clear what exactly makes people forget this fact, or to consider their own mental states as more complex than those of others.

In this regard, further work could and should attempt to better understand the role of motivated and non-motivated biases and the degree of their influence in the kind of effects that we have been describing. There is a long debate about the role of each kind of bias (e.g. Brown, 2012; Chambers & Windschitl, 2004; Sherman & Sherman, 1999). Chambers and Windschitl cleverly point out that results many times regarded as supporting a motivated

reasoning bias could very well be non-motivated (cognitive). To put it simply, as the authors explain, motivational accounts suggest that the biases that we so often observe in comparative judgments are in the service of some self-relevant goal or motive. On the other hand, non-motivated accounts consider that cognitive mechanisms and the characteristics of the task (e.g., such as the inherent actor-observer perspective that we mentioned before) are sufficient for the biases to occur. Thus, non-motivated accounts do not assume that comparative biases are consequences of there being a purpose behind people's answers, only that the answers themselves are a consequence of cognitive mechanisms and the aspects of the task itself. The former posits there being an unconscious or conscious intent, the latter does not suppose there being intent at all.

Arguably, neither can account for everything on their own, nor is it out of the possibilities that both might serve a role in self-other differences. Chambers and Windschitl (2004) conclude that "there is only limited evidence that proximal motivational concerns (feeling bad, anxious, threatened within a particular setting) trigger motivationally biased responding in comparative judgments. Whether more distal motivational concerns are influential in comparative biases might be another matter" (p. 817).

Motivation does not seem to possibly manifest itself and reach its goals without cognition, and cognition without motivation sounds very much like a machine working without purpose, objectives or goals, further implying that there can be potential motivations behind cognitive function, as well as "cognitions" giving rise to motives.

Amongst all the possible interactions between motivation and cognition, the problem presents itself a lot like a chicken-or-egg kind of question. Only here, just like the deadly-living or living-dead cat of Schrödinger, both forms of biases are as much the egg as they are the chicken. Both motivational accounts and non-motivated accounts have served each other over time in research, both informing the other's considerations. It is no longer time for

confrontation over each one is in fact the only one. Making this, like Sherman and Sherman (1999) put it, not an "unsolvable" debate but rather a debate not so useful to solve. The authors suggest moving on from this and working hard on "figuring out exactly when, why, and how motives and cognitions interact" (p. 68).

Moreover, together with this, another suggestion is to better understand what lies beyond the broad scope of the introspective process that we base our research upon. Not all the information that one might come about through the process of introspection is of the same nature. It is also not clear that each part of the whole sum of the information holds the same value compared to the remaining pieces. During the process of introspection, some thoughts, feelings and memories could be more important than others to explain and understand different phenomena. Do we expect *love* to be comparable in its extent to be explained with the ability to appreciate a beautiful landscape? Do we expect them to mean the same to someone's personal narrative and personal identity? We understand different phenomena to be perceived differently in their scientific explicability, but are they fully unexplainable, or does the extent to which they are unexplainable depend on the quality and quantity of said unexplainable parts?

It is possible for introspection to vary from a descriptive end of a continuum to a more analytic end (Ellis, 1991). Meaning, from where individuals describe their thoughts and feelings (descriptive end) to where they examine what they felt and why they felt it (analytic end). Maybe people believe that their inner self is simply indescribable (descriptive limitations). Maybe their expectations of what science might say about the reasons they feel what they feel do not match the reasons they have to feel what they feel (analytic limitations). Maybe some phenomena draw the line on these limitations differently than others.

In future research, wording the questions about each phenomenon in different ways might help to further solve this problem. The wording of questions has shown to influence

people's answers (e.g., Holleman, 1999; Loftus & Zanni, 1975; Petrinovich & Neill, 1996), and thus it may be possible to guide information processing into different parts of the same phenomenon by wording questions differently (e.g., *how*, *why*, or *what*). Maybe for some phenomena people accept science to be able to explain what the phenomenon is, but not exactly why it is what it is. An anecdotal example of this could be the problem of the meaning of life, that is, we may accept for science know what is to be alive, how are we alive, but maybe not so much why are we alive.

Thus, work should be done in figuring out exactly why and what are the perceived insufficiencies of science, now that we know that it is in fact perceived as such, as well as in understanding more clearly the contents emerging from the introspective process, so that we can better define and characterize introspection (e.g., Boring, 1953).

The implications of these results include the following: First, the skepticism concerning what science can explain about oneself should work as a barrier for people to accept certain treatments and follow certain science-based recommendations for themselves. Moreover, the double standard that people have for the limits of science for themselves vs. others should also contribute to a choice-advice asymmetry, whereby the choices that people make for themselves differ from the choices that they advise others to make (Atanasov, 2015; Dana & Cain, 2015). This is particularly relevant considering that decision-making in domains such as health can be delegated to surrogates (Buchanan & Brock, 1989; Fagerlin, Ditto, Danks & Houts, 2001). Thus, people may suggest treatments and interventions for others that they would not choose for themselves (Schroeder et al., 2017).

Moreover, if science does not understand why people dismiss its findings, how can it learn about communicating them better? What happens one day, beholding an urgent problem, if the good solution science has to offer is simply thrown aside in favor of the intuitions of those in power making decisions? This is particularly a problem for Psychology,

since people have shown to find the same scientific ideas to be better or worse depending on their scientific source (e.g., neuroscience vs. psychology; Fernandez-Duque, 2017; Weisberg, Keil, Goodstein, Rawson, & Gray, 2008), suggesting a differentiated resistance to these scientific ideas (for a developmental account of this see Bloom, & Weisberg, 2007) and a possible resistance to explanations about the mind and the endeavors of the heart compared to explanations about the physical body.

People have been "awarded" for trying to do research on things such as love, the so called "unneeded" research (see Hatfield, 2006; for an historical anecdote on this), which, considering what we have been discussing, does not seem as strange as one might have thought when learning about this story for the first time.

References

Aarssen, L. (2018). Meet Homo absurdus--the only creature that refuses to be what it is. *Ideas in Ecology and Evolution*, *11*.

Alexander, R. D. (1990). How did humans evolve? Reflections on the uniquely unique species. *Museum of Zoology (Special Publication No. 1)*. Ann Arbor, MI: The University of Michigan.

Alicke, M. D. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology*, *49*, 1621-1630.

Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. *The Self in Social Judgment, 1*, 85-106.

Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology*, *68*(5), 804.

Allison, S. T., Messick, D. M., & Goethals, G. R. (1989). On being better but not smarter than others: The Muhammad Ali effect. *Social Cognition*, *7*, 275-296.

Astington, J. W. (1993). *The child's discovery of the mind* (Vol. 31). Harvard University Press.

Atanasov, P. D. (2015). Risk preferences in choices for self and others: Meta analysis and research directions. *Available at SSRN*.

Balcetis, E., & Dunning, D. (2013). Considering the situation: Why people are better social psychologists than self-psychologists. *Self and Identity*, *12*(1), 1-15.

Becker, E. (1971). *The Birth and Death of Meaning: An Interdisciplinary Perspective on the Problem of Man. 2d Ed*. Free Press.

Bloom, P., & Weisberg, D. S. (2007). Childhood origins of adult resistance to science. *Science*, *316*(5827), 996-997.

Boring, E. G. (1953). A history of introspection. *Psychological Bulletin*, *50*(3), 169.

Brown, J. D. (2012). Understanding the better than average effect: Motives (still) matter. *Personality and Social Psychology Bulletin*, *38*(2), 209-219.

Buchanan, A. E., & Brock, D. W. (1989). *Deciding for others: The ethics of surrogate decision making*. Cambridge University Press.

Camus, A. (1956). *The Rebel*. New York: Knopf.

Chambers, J. R., & Windschitl, P. D. (2004). Biases in social comparative judgments: the role of nonmotivated factors in above-average and comparative-optimism effects. *Psychological Bulletin*, *130*(5), 813.

Chiu, C. Y., Hong, Y. Y., & Dweck, C. S. (1997). Lay dispositionism and implicit theories of personality. *Journal of Personality and Social Psychology*, *73*(1), 19-30.

Dana, J., & Cain, D. M. (2015). Advice versus choice. *Current Opinion in Psychology*, *6*, 173-176.

Davison, W. P. (1983). The third-person effect in communication. *Public Opinion Quarterly, 47*(1), 1-15.

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5*(3), 69-106.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology, 57*, 1082-1090.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality and development.* Philadelphia: Psychology Press.

Ellis, C. (1991). Sociological introspection and emotional experience. *Symbolic Interaction, 14*(1), 23-50.

Fagerlin, A., Ditto, P. H., Danks, J. H., & Houts, R. M. (2001). Projection in surrogate decisions about life-sustaining medical treatments. *Health Psychology, 20*(3), 166-175.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Fernandez-Duque, D. (2017). Lay Theories of the Mind/Brain Relationship and the Allure of Neuroscience. In *The Science of Lay Theories* (pp. 207-227). Springer, Cham.

Flinn, M. V., Geary, D. C., & Ward, C. V. (2005). Ecological dominance, social competition, and coalitionary arms races: Why humans evolved extraordinary intelligence. *Evolution and Human Behavior, 26*(1), 10-46.

Furnham, A. (1988). *International series in experimental social psychology, Vol. 17. Lay theories: Everyday understanding of problems in the social sciences.* Elmsford, NY, US: Pergamon Press.

Furnham, A., & Cheng, H. (2000). Lay theories of happiness. *Journal of Happiness Studies, 1*(2), 227-246.

Furnham, A., & Lowick, V. (1984). Lay theories of the causes of alcoholism. *British Journal of Medical Psychology, 57*(4), 319-332.

Gottlieb, S. (2018). *Can we explain the human mind and experience? Perceptions of the limits of science* (Doctoral dissertation, UC Berkeley).

Gottlieb, S., & Lombrozo, T. (2017). Can science explain the human mind? Intuitive judgments about the limits of science. *Psychological Science, 29*(1), 121-130.

Hatfield, E. (2006). The Golden Fleece Award: Love's labours almost lost. *Observer*, *19*(6), 16–17.

Heath, C. (1999). On the social psychology of agency relationships: Lay theories of motivation overemphasize extrinsic incentives. *Organizational Behavior and Human Decision Processes, 78*(1), 25-62.

Helzer, E. G., & Dunning, D. (2012). Why and when peer prediction is superior to self-prediction: The weight given to future aspiration versus past achievement. *Journal of Personality and Social Psychology, 103*(1), 38-53.

Holleman, B. (1999). Wording effects in survey research using meta-analysis to explain the forbid/allow asymmetry. *Journal of Quantitative Linguistics*, *6*(1), 29-40.

Humphrey, N. K. (1976). The social function of intellect. In *Growing Points in Ethology* (pp. 303-317). Cambridge University Press.

Johansson, P., Hall, L., Sikström, S. & Olsson, A. (2005). Failure to detect mismatches between intention and outcome in a simple decision task. *Science, 310,* 116-119

Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the cause of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79–94). Morristown, NJ: General Learning Press.

Knight, M. T., Furnham, A. F., & Lester, D. (2000). Lay theories of suicide. *Personality and Individual Differences*, *29*(3), 453-457.

Kruger, J., & Gilovich, T. (1999). " Naive cynicism" in everyday theories of responsibility assessment: On biased assumptions of bias. *Journal of Personality and Social Psychology*, *76*(5), 743-753.

Labroo, A. A., & Mukhopadhyay, A. (2009). Lay theories of emotion transience and the search for happiness: A fresh perspective on affect regulation. *Journal of Consumer Research*, *36*(2), 242-254.

Lévy-Leboyer, C. (1988). Success and failure in applying psychology. *American Psychologist*, *43*(10), 779-785.

Loftus, E. F., & Zanni, G. (1975). Eyewitness testimony: The influence of the wording of a question. *Bulletin of the Psychonomic Society*, *5*(1), 86-88.

Malle, B. F. (2006). The actor-observer asymmetry in attribution: A (surprising) meta-analysis. *Psychological bulletin*, *132*(6), 895.

Markus, H. (1983). Self-knowledge: An expanded view. *Journal of personality*, *51*(3), 543-565.

Messick, D. M., Bloom, S., Boldizar, J. P., & Samuelson, C. D. (1985). Why we are fairer than others. *Journal of Experimental Social Psychology*, *21*, 480-500.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological review*, *84*(3), 231.

Ong, D. C., Zaki, J., & Goodman, N. D. (2015). Affective cognition: Exploring lay theories of emotion. *Cognition*, *143*, 141-162.

Petrinovich, L., & O'Neill, P. (1996). Influence of wording and framing effects on moral intuitions. *Ethology and Sociobiology*, *17*(3), 145-171.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral and Brain Sciences*, *1*(4), 515-526.

Petitmengin, C., Remillieux, A., Cahour, B., & Carter-Thomas, S. (2013). A gap in Nisbett and Wilson's findings? A first-person access to our cognitive processes. *Consciousness and Cognition*, *22*(2), 654-669.

Preston, C. E., & Harris, S. (1965). Psychology of drivers in traffic accidents. *Journal of Applied Psychology*, *49*(4), 284.

Pronin, E. (2008). How we see ourselves and how we see others. *Science*, *320*(5880), 1177-1180.

Pronin, E. (2009). The introspection illusion. *Advances in Experimental Social Psychology*, *41*, 1-67.

Pronin, E., Berger, J., & Molouki, S. (2007). Alone in a crowd of sheep: Asymmetric perceptions of conformity and their roots in an introspection illusion. *Journal of Personality and Social Psychology, 92*(4), 585-595.

Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: divergent perceptions of bias in self versus others. *Psychological review*, *111*(3), 781.

Pronin, E., & Kugler, M. B. (2007). Valuing thoughts, ignoring behavior: The introspection illusion as a source of the bias blind spot. *Journal of Experimental Social Psychology*, *43*(4), 565-578.

Pronin, E., & Kugler, M. B. (2010). People believe they have more free will than others. *Proceedings of the National Academy of Sciences*, *107*, 22469-22474

Pronin, E., Kruger, J., Savtisky, K., & Ross, L. (2001). You don't know me, but I know you: The illusion of asymmetric insight. *Journal of Personality and Social Psychology, 81*(4), 639-656.

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin, 28*(3), 369-381.

Schroeder, J., Waytz, A., & Epley, N. (2017). Endorsing help for others that you oppose for yourself: Mind perception alters the perceived effectiveness of paternalism. *Journal of Experimental Psychology: General, 146*(8), 1106-1125.

Sedikides, C., Meek, R., Alicke, M. D., & Taylor, S. (2014). Behind bars but above the bar: Prisoners consider themselves more prosocial than non-prisoners. *British Journal of Social Psychology, 53*(2), 396-403.

Sherman, S. J., & Sherman, J. W. (1999). Bring the troops back home: armistice between motivation and cognition. *Psychological Inquiry, 10*(1), 65-68.

Smith, E. R., & Miller, F. D. (1978). Limits on perception of cognitive processes: A reply to Nisbett and Wilson. *Psychological Review, 85*, 355-362.

Sommers, T., & Rosenberg, A. (2003). Darwin's nihilistic idea: evolution and the meaninglessness of life. *Biology and Philosophy, 18*(5), 653-668.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers?. *Acta Psychologica*, *47*(2), 143-148.

Tamir, M., John, O. P., Srivastava, S., & Gross, J. J. (2007). Implicit theories of emotion: Affective and social outcomes across a major life transition. *Journal of Personality and Social Psychology, 92*(4), 731-744.

Waytz, A., Schroeder, J., & Epley, N. (2014). The lesser minds problem. *Humanness and Dehumanization*, 49-67.

Weisberg, D. S., Keil, F. C., Goodstein, J., Rawson, E., & Gray, J. R. (2008). The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience, 20*(3), 470-477.

Williams, E. F., & Gilovich, T. (2008). Do people really believe they are above average?. *Journal of Experimental Social Psychology*, *44*(4), 1121-1128.

Wilson, T. D., & Gilbert, D. T. (2003). Affective forecasting. *Advances in Experimental Social Psychology, 35*(35), 345-411.