COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 45 (2021) e13034

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13034

Neurodualism: People Assume that the Brain Affects the Mind more than the Mind Affects the Brain

Jussi Valtonen,[a] Woo-kyoung Ahn,[b] Andrei Cimpian[a]

[a]*Department of Psychology, New York University*
[b]*Department of Psychology, Yale University*

Received 22 December 2020; received in revised form 6 June 2021; accepted 9 July 2021

Abstract

People commonly think of the mind and the brain as distinct entities that interact, a view known as dualism. At the same time, the public widely acknowledges that science attributes all mental phenomena to the workings of a material brain, a view at odds with dualism. How do people reconcile these conflicting perspectives? We propose that people distort claims about the brain from the wider culture to fit their dualist belief that minds and brains are distinct, interacting entities: Exposure to cultural discourse about the brain as the physical basis for the mind prompts people to posit that mind–brain interactions are *asymmetric*, such that the brain is able to affect the mind more than vice versa. We term this hybrid intuitive theory *neurodualism*. Five studies involving both thought experiments and naturalistic scenarios provided evidence of neurodualism among laypeople and, to some extent, even practicing psychotherapists. For example, lay participants reported that "a change in a person's brain" is accompanied by "a change in the person's mind" more often than vice versa. Similarly, when asked to imagine that "future scientists were able to alter exactly 25% of a person's brain," participants reported larger corresponding changes in the person's mind than in the opposite direction. Participants also showed a similarly asymmetric pattern favoring the brain over the mind in naturalistic scenarios. By uncovering

---

We are grateful to Sarah Hulsmann, Angelica Mehta, Jasmine Miller, Liam North, and Molly Tallberg for their help in conducting this research. This research was supported by postdoctoral research grants for Jussi Valtonen from the Finnish Cultural Foundation and the Kone Foundation. The data from all studies reported in this manuscript can be accessed on the Open Science Framework: https://osf.io/3sknp/?view_only=d33490ae62d04d8d82af49fd5ccaad4f

Correspondence should be sent to Jussi Valtonen. He is now at the Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, P.O. Box 21, 00014 Helsinki, Finland. E-mail: jussi.valtonen@helsinki.fi

people's intuitive theories of the mind–brain relation, the results provide insights into societal phenomena such as the allure of neuroscience and common misperceptions of mental health treatments.

Keywords: Causality; Intuitive theories; Mental disorders; Mind–body dualism; Reasoning

1. Neurodualism: People assume that the brain affects the mind more than the mind affects the brain

Mind–body dualism refers to the view that the mental and the physical both exist and that neither can be assimilated to the other (for a philosophical review, see Robinson, 2017). A common claim among cognitive theorists is that humans are intuitive dualists, perceiving the mind and the body as distinct from each other (e.g., Bloom, 2004). While philosophers discern among several different dualist positions, lay intuitions are typically thought to consist of a pair of simple tenets: that the brain is somehow different from the mind and that the two can interact with each other. Churchland (2013) refers to this view as *popular dualism*:

> This is the theory that a person is literally 'a ghost in a machine', where the machine is the human body, and the ghost is a spiritual substance, quite unlike physical matter in its internal constitution, but fully possessed of spatial properties even so. In particular, minds are commonly held to be *inside* the bodies they control: inside the head, on most views, in intimate contact with the brain. (p. 15)

Dualistic beliefs of this sort seem common among laypeople: A survey of 160 American undergraduate students found that 50% agreed with the statement that the mind is a special form of energy that is currently unknown to science but that is in contact with the brain, and 49% of participants also subscribed to the popular dualist view that the mind interacts with the brain to determine behavior (Stanovich, 1989). An even higher proportion of Americans, 75–85%, believe in life after the death of the material body (Greeley & Hout, 1999), a belief that entails dualism. Surveys in European countries have found similar levels of endorsement of popular dualism (e.g., more than 60% of the Belgian and Scottish participants queried disagreed with the claim that the mind is fundamentally physical; Demertzi et al., 2009), but—notably—popular dualism is not restricted to industrialized, Western nations (Bering, 2006; Chudek, McNamara, Birch, Bloom, & Henrich, 2018). For example, a recent study conducted in the United States, Ghana, Thailand, China, and Vanuatu found that the mind–body distinction was common across these cultures and emerged in childhood (Weisman et al., in press). While these findings stop short of demonstrating that popular dualism is a universal human tendency, they nevertheless point to its prevalence across a range of cultures (see also Chudek et al., 2018; Hodge, 2008; Slingerland & Chudek, 2011).

While dualistic beliefs may be common, it is hardly a secret from the general public that scientific knowledge attributes mental phenomena to the workings of a material brain. How, then, do people reconcile these two discrepant views—the intuition of the mind and the body as separate entities on the one hand, and of the brain as the material organ that enables the mind on the other hand? Apart from the general recognition that people tend to be intuitive (popular) dualists, surprisingly little is known about how people resolve this tension.

To make sense of the world, people rely on intuitive theories—abstract domain-specific cognitive structures that guide explanations and predictions (Ahn, 1998; Gerstenberg & Tenenbaum, 2017; Gopnik & Wellman, 1992). A better understanding of lay intuitive theories about the mind and the brain is important not only for insights into human cognition, but also because intuitive theories contain causal beliefs, and causal beliefs affect behavior. Dualistic biases about the mind and the body have been shown to affect, for example, health-related behaviors (Forstmann, Burgmer, & Mussweiler, 2012), expectations of mental-health treatment outcomes (Ahn, Kim, & Lebowitz, 2017; Deacon & Baird, 2009; Kemp, Lickel, & Deacon, 2014; Lebowitz & Ahn, 2015), mental illness stigma (Haslam & Kvaale, 2015), the clinical judgments of mental health professionals (Miresco & Kirmayer, 2006), and, some authors argue, implicitly even ways in which entire societal institutions, such as health care and the criminal justice system, are structured (Greene, 2011). Both laypeople and mental health professionals believe, for example, that a biological treatment is more effective for mental health problems with a biological basis, a belief that encourages increased use of psychiatric drugs and decreased utilization of psychotherapy as the neuroscience of mental disorders progresses (Ahn, Proctor, & Flanagan, 2009; Ahn et al., 2017). A better understanding of intuitive theories about the mind and the brain can thus shed new light on consequential decisions and behaviors.

In this paper, we propose that people tend to (1) view the mind and the brain as distinct, interacting entities, a notion consistent with dualism, but also (2) extrapolate the notion that the brain is the biological source of the mind—a notion incompatible with popular dualism—into a belief that the brain is more likely to affect the mind than the mind is to affect the brain. We will first briefly review the cognitive literature motivating this proposal and then present evidence from five studies testing it.

1.1. How do people think about minds and brains?

Our goal here is to identify and test the intuitive mind–brain theory that emerges from our ingrained dualist tendencies on the one hand and from the nondualist assumptions acquired through education, media, and exposure to the broader culture on the other hand. In this section, we briefly review evidence for popular dualism in reasoning about the mind and the brain/body, as well as for the other beliefs that—we argue—complicate simple dualism.

Mind–body dualism emerges early in development (Bloom, 2004; Wellman & Johnson, 2008; Weisman et al., in press). Children use different cognitive frameworks to think about the biological and the psychological (Bloom, 2007; Wellman & Gelman, 1992)—foundational knowledge systems that shape children's causal attributions and perceptions of causal forces involving minds and brains (Carey, 1985; Inagaki & Hatano, 1993). Dualistic biases shape reasoning among adults as well (Ahn et al., 2009; Forstmann et al., 2012; Miresco & Kirmayer, 2006; Proctor, 2008). Forstmann and Burgmer (2015), for example, showed that adults believe that a living creature would retain a greater number of physical than mental properties if they were physically duplicated, consistent with the notion that people are intuitive dualists (see also Horne & Cimpian, 2019).

Dualistic intuitions are ingrained in human thought also in the sense that education often does not eradicate them. For example, trained mental health professionals perceive the biological and psychological domains as surprisingly separate, despite scientific evidence suggesting that this view is not accurate (Ahn et al., 2009). In the words of Miresco and Kirmayer (2006), "mental health professionals continue to employ a mind–brain dichotomy when reasoning about clinical cases" (p. 913). Even neuroscientists commonly discuss the mind and the brain in ways that arguably reveal underlying dualistic beliefs (Mudrik & Maoz, 2014), suggesting that such intuitions are deep-seated.

Although dualistic intuitions are prevalent, people hardly perceive the brain as irrelevant for behavior. The scientific community widely endorses and communicates the view that everything in the world is fundamentally material or physical, a view known as materialism or physicalism.[1] The general population is increasingly confronted with the view of the physical brain as the seat of all cognition, emotion, and behavior—a notion also widely reflected in popular culture (e.g., Beck, 2010). As Greene (2011, p. 265) puts it, "These days, even the most ardent dualists recognize that we have brains and that our brains must do *something*." Consistent with this claim, the majority of respondents in a survey of more than 1800 Belgian adults agreed with the statement that the mind is *not* separable from the brain (Demertzi et al., 2009), a statement at odds with dualism by definition (see also Stanovich, 1989). In light of these considerations, one may wonder whether lay theories of the mind and the brain are adequately described by the tenets of popular dualism (Churchland, 2013).

How, then, do people *actually* reason about what happens in the mind and in the brain? How do people reconcile the two conflicting beliefs—the dualistic intuition that the mind and the brain are separate on the one hand, and the notion of the brain enabling, or even being identical to, the mind on the other hand?

1.2. Neurodualism: A proposed intuitive theory of mind–brain interactions

The claim that everything in the world is fundamentally physical contradicts the entrenched dualist belief that the brain and the mind are distinct: Intuitively, our thoughts and desires do not *seem* physical to us. Thus, physicalism needs to be altered before it can be incorporated into intuitive theories of minds and brains. Rather than adopting physicalism per se, people distort scientific discourse about the brain into a simpler belief: namely, that the brain is *particularly important* for how our minds work. This belief—unlike the original physicalist claim—can be incorporated into intuitive theories of minds and brains because popular dualism already allows that the brain and the mind, although distinct, can interact with each other.

There are numerous situations in everyday life that reinforce this simpler belief by reminding us of how important the brain is for our mental lives. For example, drinking alcohol or coffee produces well-known psychological effects by intervening on brain function. Some of us are also forced to bear witness to how a traumatic brain injury or a progressive brain disease can incapacitate a loved one's mental faculties. Intuitively, it is easier to interpret such instances as examples of the brain *affecting* the mind rather than as evidence of everything in the world being fundamentally physical. In fact, this notion that the brain affects—but is distinct from—the mind is so intuitive and so common that even neuroscientists sometimes let it slip into their work. That is, neuroscientists sometimes refer to the brain and the mind as if they were two separate agents with, for example, brain activity temporally preceding the corresponding mental states (Mudrik & Maoz, 2014).

In contrast to the everyday examples which intuitively suggest to us that the brain affects our mental lives, we rarely encounter situations that can be interpreted as evidence of our mental lives affecting our physical brains: We never see our thoughts, intentions, or emotions manifested in our brains. Only very recently has the lay public been exposed to brain imaging studies in the popular media, but these studies are far removed from everyday experience. For these reasons, people may be reluctant to believe that our mental lives could affect our physical brains as much as our brains affect our minds. Princess Elisabeth of Bohemia expressed this reluctance elegantly in her famous 1643 letter to Descartes, in which she asked the philosopher to explain "how the mind of a human being can determine the bodily spirits (i.e., the fluids in the nerves, muscles, etc.) in producing voluntary actions, being only a thinking substance" (quoted in Kim, 2010, p. 243). Conceivably, this same difficulty could bias our contemporaries' intuitions in favor of seeing the brain as a more powerful causal agent relative to the mind.

To summarize, we propose the following: Intuitive theories of minds and brains incorporate from broader scientific discourse a distorted form of the notion that the brain is the physical seat of all mental life: namely, that *the brain generally has more power over the mind than the mind does over the brain*. In this way, intuitive theories can retain the core dualist belief that the mind and the brain are distinct, but the physicalist claim about the brain as the physical basis for all events in the mind can also be accommodated—in this distorted form. The end result is a hybrid intuitive theory of the mind–brain relation which posits that (1) the brain and the mind are distinct and interacting (the tenets of popular dualism) but also that (2) mind–brain interactions have an asymmetric character, such that the brain is able to affect the mind more than the mind is able to affect the brain. We term this proposed intuitive theory *neurodualism* to highlight the assumed causal asymmetry favoring the brain.

Such an intuitive theory for reasoning about minds and brains has, to our knowledge, been neither previously proposed nor empirically investigated. However, the developmental literature is consistent with the notion of a hybrid intuitive theory, broadly conceived: Children learn to incorporate physicalist notions into their beliefs already at an early age, but they do so without surrendering their more fundamentally dualistic worldviews. For example, by first grade, children learn that a person needs a brain for thinking (Johnson & Wellman, 1982), but they do not believe that all their mental activity is dependent on the brain (Bloom, 2004; Lillard, 1996). Although our proposal is different in important respects, it makes the similar claim that some aspects of dualism are retained even as physicalist (or physicalism-inspired) beliefs are acquired.

1.3. The present research

Across five experiments, we tested whether laypeople's and mental health professionals' intuitions about the effects of changes in the mind and in the brain are consistent with neurodualism. We used a variety of question types to triangulate on participants' intuitive beliefs about the mind and the brain, including (1) abstract thought experiments that asked directly about the effects of changes in the mind and the brain, and (2) concrete naturalistic scenarios that embedded the questions about changes in the mind and the brain in the context of ostensibly real-life individuals undergoing psychotherapy or taking psychiatric medications. These two types of tasks have complementary strengths and weaknesses. On the one hand, abstract thought experiments are less likely to activate extraneous beliefs about real-life details and may thus provide a cleaner measure of people's intuitive theories. In our studies, we asked participants to imagine that future scientists were able to alter the mind and the brain directly, and asked how these alterations would affect the other. On the other hand, the greater detail in naturalistic scenarios makes them more ecologically valid, and thus arguably better at capturing how people reason in real-life contexts. To manipulate mind-versus-brain interventions in naturalistic scenarios, we asked participants to consider individuals who suffer from various problems and then undergo interventions with which most laypeople are reasonably familiar: psychotherapy (i.e., intervening on the mind; Studies 1–5) or psychiatric medications (i.e., intervening on the brain; Studies 3–5).[2]

In Studies 1–4, we investigated these intuitive beliefs in lay participants. Study 5 extended the focus from lay participants to licensed psychotherapists in the United States. All studies were approved by the ethics board of the institution in which the research was carried out. Thus, Studies 1–4 were approved by New York University's Institutional Review Board and Study 5 by the Yale University Institutional Review Board.

2. Study 1: The locus-of-change and hypothetical-alterations tasks

In Study 1, we began by asking participants (1) whether they believed psychotherapy, a real-life intervention acting on the mind, has a greater effect on the mind or on the brain, and (2) how much they believed hypothetical interventions directly altering the mind or the brain would affect the other.

Different patterns of results are predicted depending on how participants reason about the mind and the brain. We go on to contrast the predictions that follow from (1) popular dualism, folk versions of (2) reductive and (3) nonreductive physicalism, and (4) neurodualism. If participants endorse *popular dualism*, they will have no reason to weigh the causal power of the brain differently from that of the mind. The same can be said if participants endorse a folk version of physicalism. For instance, if participants endorse *a strict, reductive form of physicalism* such as identity theory (Smart, 1959), which assumes that mental events are identical to brain events, then they should think it is impossible to intervene on mental events and not intervene on brain events equally.[3] If intervening on the mind has any effects at all on psychological processes, these effects must be ones that affect brain processes to an equal extent. As yet another possibility, participants could use *a more flexible, nonreductive form of physicalism*, such as supervenience (Kim, 1998), as a basis for their intuitions. Supervenience is compatible with asymmetric relations between brains and minds: For example, if two people have different brain states at one point in time, they may or may not have different (supervening) mental states. Supervenience, however, does not allow changes in mental states that are not also changes in brain states. Therefore, if the starting point for lay intuitive theories is something like supervenience, participants should be more likely to accept the possibility of changes in brain states not resulting in changes in mental states than vice versa—the exact opposite of the prediction from neurodualism. Finally, if participants adopt *a neurodualist intuitive theory*, they should be *less* likely to report that a change in the mind results in a change in the brain than the other way around. This prediction is unique to neurodualism—it does not follow either from popular dualism or from various forms of physicalism, including identity theory and supervenience.

Study 1 tested this prediction in three different ways, using different wordings. Arguably, the simplest and most intuitive way to ask participants about changes in the mind and brain is to do exactly that. In Study 1A, we therefore asked about changes "in the mind" and "in the brain." A potential concern about this wording, however, is that participants might interpret these questions in ways other than we intended. For example, when asked whether psychotherapy "changes the brain," participants may think, understandably, that psychotherapy will not fundamentally alter the brain's gross anatomical structure. We, therefore, included two additional conditions using terminology that refers clearly to transient and malleable events that cannot be confused with gross anatomical brain structures. In Study 1B, we asked participants about changes in "brain processes" versus "mental processes," and in Study 1C, about changes in "brain states" versus "mental states." If participants respond in a consistent way across Studies 1A–1C despite the changes in terminology, the results will provide stronger evidence about the underlying intuitive beliefs than if the findings hinge on a particular wording or interpretation.

2.1. Method

2.1.1. Participants

Participants were recruited online through Amazon's Mechanical Turk (MTurk) platform and paid $0.50 for participating. Only MTurk workers within the United States were allowed to participate. All participants gave informed consent online before starting the survey. Workers who participated in one of the three studies were blocked from participating in the others. We recruited a total of 177 participants (68 females, 95 males, 14 chose not to disclose their gender identity; mean age = 34.2 years; age range = 21–75 years; modal education = 4-year college degree). Of these, 46 participants were excluded for failing an attention check (see Materials and Procedure). After exclusions, the sample sizes were 43, 43, and 45 for Studies 1A, 1B, and 1C (48.8%, 53.5%, and 60.0% male), respectively. A sensitivity power analysis (power = 80%, $\alpha = .05$, two-tailed tests) suggested that these studies were adequately powered to detect a within-subject effect as small as Cohen's $d_z = 0.44$.

2.1.2. Materials and procedure

We included two measures of participants' intuitive theories of mind–brain interactions. Table 1 includes a summary of the measures used across Studies 1–5.

2.1.2.1. Locus of change in response to psychotherapy. To investigate the perceived effects of a psychological intervention, we asked participants to imagine different scenarios

Table 1
The measures used in Studies 1–5, with sample items

		Sample Items	
Task	Studies	Intervention on the **brain**	Intervention on the **mind**
Locus of change	1		A person has difficulties in their personal life. They see a therapist for 3 months. Would this treatment cause a greater change in the person's [mind; Study 1A] [mental processes; Study 1B] [mental states; Study 1C] or in their [brain; Study 1A] [brain processes; Study 1B] [brain states; Study 1C]?
	2		A person has difficulties in their personal life. They see a therapist for 3 months. How much of a change would this treatment cause in the person's mind? And how much of a change would this treatment cause in the person's brain?
	3	A person has difficulties in their personal life. They **take medication** for 3 months and completely overcome their difficulties. How much of a change would this treatment cause in the person's mind? And how much of a change would this treatment cause in the person's brain?	A person has difficulties in their personal life. They **see a therapist** for 3 months and completely overcome their difficulties. How much of a change would this treatment cause in the person's mind? And how much of a change would this treatment cause in the person's brain?
	4–5	Terry is a 28-year-old woman who is seeking treatment because she has felt deeply sad for the past 4 weeks. She reports that she has lost interest in activities that normally bring her pleasure, such as cooking and going to movies. She states that ever since the current period of sadness began, she has had difficulty with her memory and concentration. She has canceled plans with friends several times because she felt too tired to go out. She reports that her outlook on life has become quite pessimistic; she describes feeling hopeless about the future and perceiving herself to be totally worthless. She has reportedly experienced similar symptoms (a low mood that lasted several weeks) on three previous occasions in her life. Terry **takes medication** for 12 months	Terry is a 28-year-old woman who is seeking treatment because she has felt deeply sad for the past 4 weeks. She reports that she has lost interest in activities that normally bring her pleasure, such as cooking and going to movies. She states that ever since the current period of sadness began, she has had difficulty with her memory and concentration. She has canceled plans with friends several times because she felt too tired to go out. She reports that her outlook on life has become quite pessimistic; she describes feeling hopeless about the future and perceiving herself to be totally worthless. She has reportedly experienced similar symptoms (a low mood that lasted several weeks) on three previous occasions in her life. Terry **sees a therapist** for 12 months

(*Continued*)

Table 1
Continued

		Sample Items	
Task	Studies	Intervention on the **brain**	Intervention on the **mind**
		and her symptoms are now under control. How much of a change did this treatment cause in Terry's mental processes? And how much of a change did this treatment cause in Terry's brain processes?	and her symptoms are now under control. How much of a change did this treatment cause in Terry's mental processes? And how much of a change did this treatment cause in Terry's brain processes?
Hypothetical alterations	1–5	Imagine that future scientists were able to alter exactly 25% of a person's [**brain**; Studies 1A, 2–5] [**brain processes**; Study 1B] [**brain states**; Study 1C]. How much of a change would this cause in their [**mind**; Studies 1A, 2–5] [**mental processes**; Study 1B] [**mental states**; Study 1C]?	Imagine that future scientists were able to alter exactly 25% of a person's [**mind**; Studies 1A, 2–5] [**mental processes**; Study 1B] [**mental states**; Study 1C]. How much of a change would this cause in their [**brain**; Studies 1A, 2–5] [**brain processes**; Study 1B] [**brain states**; Study 1C]?
Changes in the abstract	3–5	When there is a change in a person's **brain,** how often would you say there is also a change in the person's **mind?**	When there is a change in a person's **mind,** how often would you say there is also a change in the person's **brain?**

in which a person saw a psychotherapist and the person's problems were treated. Participants then judged to what extent this treatment changed the person's mind and brain. We will refer to this measure as the *locus-of-change task*.

To assess the generalizability of the results, we administered scenarios that differed in the reason the person was seeing a therapist. In some scenarios, participants were told, "A person has difficulties in their personal [professional] life," and in others, "A person has been diagnosed with depression [anxiety]." For each scenario, we created two versions that differed in treatment duration (3 or 12 months; e.g., "They see a therapist for 3 months"). In total, there were 4 (reasons) $\times$ 2 (durations) = 8 different scenarios.

To measure the impact of psychotherapy, the participants in Study 1A were asked, "Would this treatment cause a greater change in the person's mind or in their brain?" Participants responded by clicking on a slider, on which one pole was labeled "Change exclusively in the MIND," the opposite pole "Change exclusively in the BRAIN," and the midpoint labeled "Identical change in MIND and BRAIN." The slider was initially positioned at its midpoint. The position of the labels was randomized, such that the left pole was labeled as "mind" for half of the participants and "brain" for the other half. The presentation order of the words "mind" and "brain" in the question was aligned accordingly, such that the left-pole label on the slider was always mentioned first in the question. All responses were recorded on a scale from -50 to $+50$, which indicated change exclusively in the mind and brain, respectively, with 0 indicating identical change in both (i.e., no asymmetry). In Studies 1B and 1C, the words "mind" and "brain" were replaced by "mental processes" and "brain processes" (Study 1B) and "mental states" and "brain states" (Study 1C).

2.1.2.2. Hypothetical alterations of the mind or the brain. For an alternative means of quantifying intuitions about minds and brains, we also asked the participants about the effects of hypothetically altering the mind or the brain *directly*. In Study 1A, the participants were presented with the following two scenarios: "Imagine that future scientists were able to alter exactly 25% of a person's brain. How much of a change would this cause in their mind?" And, "Imagine that future scientists were able to alter exactly 25% of a person's mind. How much of a change would this cause in their brain?" We will refer to this measure as the *hypothetical-alterations task*. The order of the two questions was randomized across participants. Participants responded by clicking on a slider whose left and right poles were labeled "0%" and "100%," respectively. The slider was initially positioned at 25%, and the participants were explicitly notified of this. In Study 1B and 1C, the words "brain" and "mind" were replaced by the phrases "brain processes" and "mental processes" (Study 1B) and "brain states" and "mental states" (Study 1C).

The locus-of-change task was always presented first, followed by the hypothetical-alterations task. We chose this fixed order because we judged that the hypothetical-alterations task (which is very direct and might make our research question apparent to participants) was more likely to interfere with responses on the locus-of-change task than vice versa.

2.1.2.3. Attention check. One item was included to assess whether participants read all items carefully before responding. The item explicitly stated it was an attention-check question, but proceeded to ask a seemingly legitimate question, after which it told participants to ignore the question and to choose the last option of six choices. Forty-six participants were excluded for failing this check.

2.1.3. Open and transparent scientific practices

The data from this and all subsequent studies can be accessed on the Open Science Framework: https://osf.io/3sknp/?view_only=d33490ae62d04d8d82af49fd5ccaad4f

2.2. Results

2.2.1. Locus-of-change task

Participants' responses to the naturalistic scenarios in the locus-of-change task were consistent with neurodualism. Regardless of question wording, the reasons for needing treatment, and the length of the treatment, participants systematically expected psychotherapy to cause greater changes in the mind than in the brain. As shown in Figure 1, the 95% confidence intervals are consistently on the negative (mind) side of zero in all versions of the scenarios. (For a description of differences as a function of reasons for needing treatment and length of the treatment, see online Appendix A.)

2.2.2. Hypothetical-alterations task

Also consistent with our prediction, participants believed that hypothetical interventions by future scientists would have asymmetric effects on the mind and the brain (see Figure 2). Regardless of the wording used, participants expected that direct alterations of the brain would

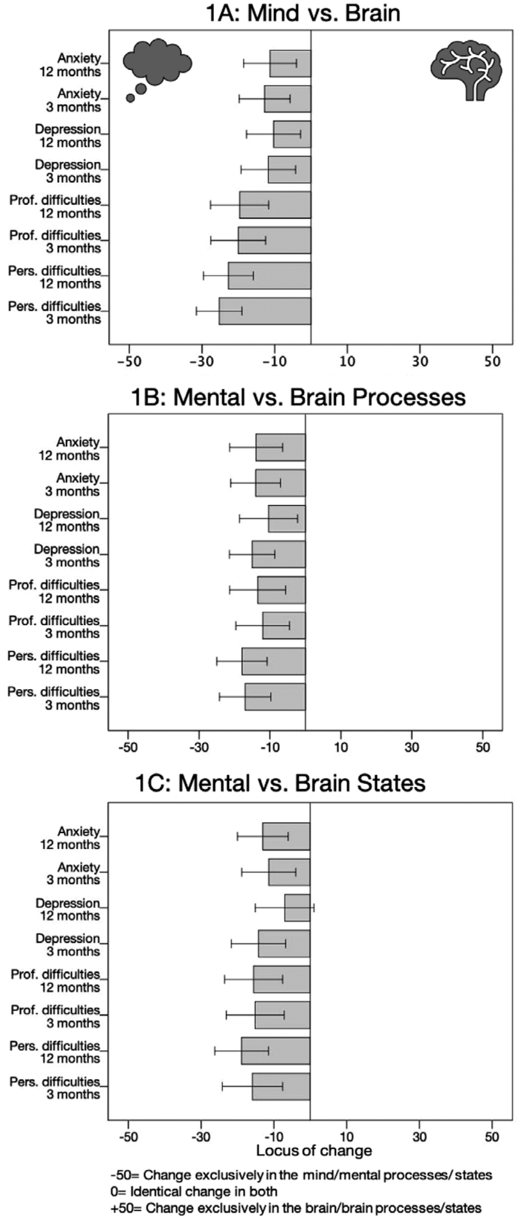

Fig 1. Participants' reports of the locus of change in psychotherapy. Participants were asked if psychotherapy would cause a greater change in the mind, the brain, or an identical change in both. The participants responded on a continuum from "exclusively in the mind" (Study 1A, top), "exclusively in mental processes" (1B, middle), and "exclusively in mental states" (1C, bottom) to "exclusively in the brain," "exclusively in brain processes," and "exclusively in brain states" (−50 to +50, respectively; 0 = "identical change in the mind and the brain"). On each trial, the participants read one of four different reasons the fictitious person was seeing a therapist (e.g., problems in personal life, depression) and were given one of two durations (3 vs. 12 months) for the treatment. The error bars depict 95% confidence intervals.

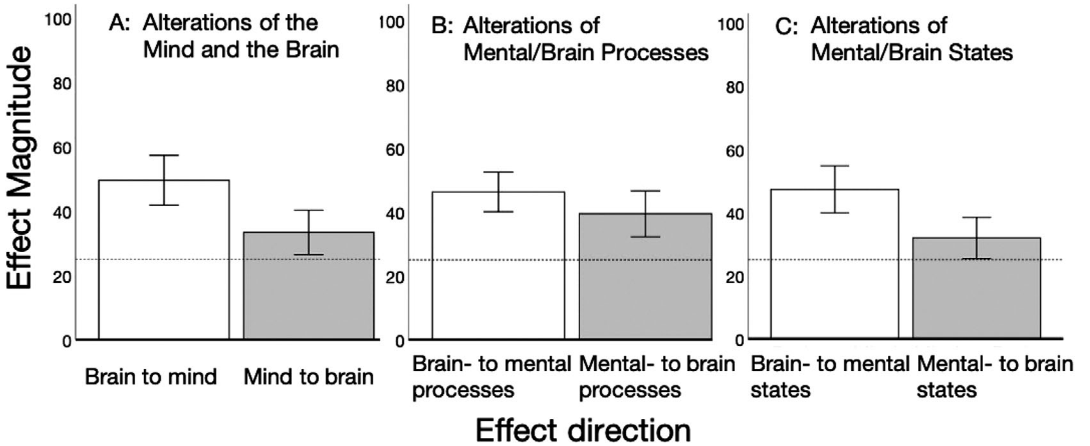

Fig 2. Participants' reports of the anticipated effects of direct alterations of the mind and the brain in Study 1. Participants were asked about direct alterations of (a) the mind and the brain, (b) mental and brain processes, and (c) mental and brain states. In all conditions, participants believed that altering the brain would change the mind more (white bars) than altering the mind (by the same amount) would change the brain (gray bars). The dotted line depicts the magnitude of the original intervention (25%). Error bars depict 95% confidence intervals.

change the mind more than direct alterations of the mind (of an identical magnitude) would change the brain.

In Study 1A, the participants believed that a 25% alteration of the brain by future scientists would change the mind by nearly 50% ($M = 49.5$, 95% CI [41.8, 57.3]), while altering the mind by the exact same magnitude would change the brain by only 33% ($M = 33.4$ [26.5, 40.3]; $F[1, 42] = 12.27$, $p = .001$, $\eta_p^2 = .226$). In Study 1B, participants expected the alteration of brain *processes* to change mental *processes* by 46%, whereas the same alteration of mental processes would only change brain processes by 39% ($M = 46.3$ [40.1, 52.5] and $M = 39.4$ [32.2, 46.6], respectively; $F[1, 42] = 5.17$, $p = .028$, $\eta_p^2 = .110$). In Study 1C, participants expected that an alteration of brain *states* would change a person's mental *states* by 47%, but only by 32% in the mind-to-brain direction ($M = 47.3$ [39.9, 54.7] and $M = 31.9$ [25.4, 38.4], respectively; $F[1, 44] = 12.99$, $p < .001$, $\eta_p^2 = .228$).

2.3. Discussion

Across two different tasks involving both naturalistic scenarios (the locus-of-change task) and thought experiments (the hypothetical-alterations task), participants reported that intervening on the mind and the brain would have asymmetric effects, consistent with neurodualism.

In naturalistic scenarios varying in the reasons why treatment was sought and in the time the treatment lasted, the participants reported systematically that psychotherapy, which is commonly understood to intervene on the mind, would change the mind more than the brain. This result is consistent with the neurodualism account, according to which intuitive theories of the mind–brain relation assume the mind is relatively powerless to change the brain. Note that it would have been perfectly possible to give a physicalist response in this study and indicate that the expected changes of psychotherapy were identical in the mind and the brain. However, participants chose not to do so, consistent with neurodualism. The predicted asymmetry was observed regardless of whether the stimulus wordings referred to "the mind and the brain," "mental processes and brain processes," or "mental states and brain states," ruling out the possibility that the findings were an artifact arising from a particular interpretation of the words "mind" and "brain."

In thought experiments, participants consistently judged that hypothetical interventions by future scientists directly altering the brain would cause greater changes in the mind than their altering the mind would cause in the brain. Again, it would have been easy for the participants to respond using the original 25% alteration as a reference point and say that the effects were 25% as well, regardless of whether the interventions occurred in the mind or brain. Yet, they chose not to do so. Instead, they always judged the effects of brain alterations on the mind to be greater than vice versa.

3. Study 2: Revising the locus-of-change task to minimize task demands

One potential limitation of the locus-of-change task used in Study 1 was that participants responded on a continuum with two opposite poles, suggesting a forced choice. Although the response scale did allow participants to place the slider at the midpoint, which was explicitly labeled as "identical change in both," the question nevertheless arises whether the asymmetry observed in Study 1 was an artifact produced by the task setup (if participants believed they had to choose between the mind and the brain). To rule out this possibility, in Study 2, we asked participants to *separately* evaluate the magnitude of the effects of psychotherapy on the mind and the brain. In addition, we sought to replicate the results of the hypothetical-alterations task from Study 1 with a new sample of participants.

3.1. Method

3.1.1. Participants

Participants were recruited through MTurk, as in Study 1. Participants who had taken part in Study 1 were prevented from responding. We recruited a total of 58 participants (21 females, 32 males, 1 other, 4 chose not reveal their gender identity; mean age = 35.0 years; range = 22–71 years; 74.1% white; modal education = 4-year college degree). Of these, 18 participants were excluded for failing an attention check (same as in Study 1).

3.1.2. Materials and procedure

All materials and procedures were the same as in Study 1A, with one exception: Participants were now asked *two* questions after each scenario in the locus-of-change task: "How much of a change would this treatment cause in the person's mind? And how much of a change would this treatment cause in the person's brain?" (see Table 1). The participants responded using two 0–100 sliding scales titled "Change in the MIND" and "Change in the

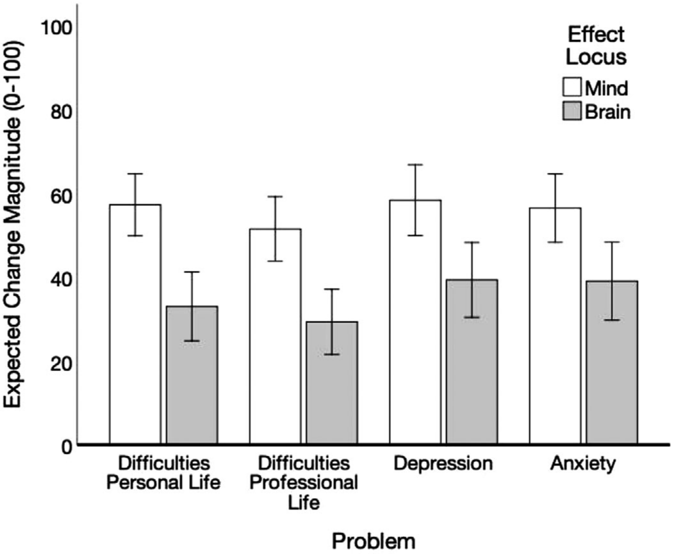

Fig 3. Participants' reports of the anticipated psychotherapy-induced changes in the mind and the brain in Study 2. Participants believed that psychotherapy would exert greater changes in the mind (white bars) than in the brain (gray bars), regardless of type of problem for which psychotherapy was sought. Participants responded using two sliders (0 = "no change" to 100 = "huge change"), one for changes in the mind and the other for changes in the brain. Error bars represent 95% confidence intervals.

BRAIN." The left and right poles of the sliders were labeled, "NO CHANGE," and "HUGE CHANGE," respectively. The spatial arrangement of the sliders was counterbalanced, such that the mind-slider was presented above the brain-slider for approximately 50% of the participants and below it for the others. The order in which the mind and the brain were mentioned in the questions corresponded with the order of the sliders, such that the domain mentioned first (mind for 50% of participants vs. brain for the others) always corresponded to the first slider.

3.2. Results

3.2.1. Locus-of-change task

Consistent with our proposal, participants reported that psychotherapy would cause greater changes in the mind than in the brain. As shown in Figure 3, the causal asymmetry was consistent across problem types ($M = 56.0$ [48.9, 63.0] and $M = 35.2$ [27.4, 43.1] for mind and brain, respectively). A 4 (problem type) × 2 (therapy duration) × 2 (locus of change) repeated-measures ANOVA showed a main effect of locus of change, $F(1, 39) = 20.04$, $p < .001$, $\eta_p^2 = .339$, but no significant two- or three-way interactions.

The participants also expected therapy to induce more change overall if it lasted 12 rather than 3 months ($M = 52.1$ [46.0, 58.2] vs. $M = 39.1$ [33.1, 45.2], respectively; $F[1, 39] = 46.92$, $p < .001$, $\eta_p^2 = .546$). Finally, we also found a main effect of problem type, $F(2.50, 97.52) = 5.48$, $p = .003$, $\eta_p^2 = .123$:[4] Participants expected therapy to induce more change in the scenarios in which the person was diagnosed with depression ($M = 48.9$ [42.3, 55.6]) or anxiety ($M = 47.8$ [40.9, 54.7]) than in ones involving difficulties in the person's professional ($M = 40.5$ [34.4, 46.6]) or personal ($M = 45.2$ [39.2, 51.2]) life. Regardless of these details, however, the anticipated primary locus of change was always in the mind.

3.2.2. Hypothetical-alterations task

The pattern of results in the hypothetical-alterations task was virtually identical to that in Study 1. Consistent with neurodualism, participants again expected a direct alteration of the brain to change the mind more than vice versa ($M = 39.4$ [31.8, 46.9] vs. $M = 28.4$ [21.2, 35.6], respectively; $F[1, 39] = 5.36$, $p = .026$, $\eta_p^2 = .121$).

3.3. Discussion

Study 2 replicated the findings from Study 1 on both the locus-of-change and the hypothetical-alterations tasks. Of note, the locus-of-change task used a new setup that allowed participants to indicate anticipated changes to the brain and the mind independently of each other—a revision to our method that was intended to investigate the possibility that the asymmetry observed in Study 1 was an artifact of the response format. Contrary to this possibility, the results showed that participants again consistently expected that a psychological intervention, psychotherapy, would cause greater changes in the mind than in the brain. These results are consistent with an underlying intuitive theory in which the causal effects of the mind on the brain are weak.

4. Study 3: Adding the changes-in-the-abstract task; comparing the effects of psychotherapy and psychiatric drugs in the locus-of-change task

In Study 3, we added a third, even more direct, measure of neurodualism: We asked participants how often they thought events in the mind are accompanied by events in the brain and vice versa. We also added an important element to the locus-of-change task. Thus far, these scenarios only investigated beliefs about the effects of intervening on the mind (psychotherapy). The results suggested that the anticipated causal effects of mind-interventions on the brain are weak. Conceivably, however, participants might believe that the effects of *any* intervention are stronger in a metaphysically proximal domain (in Studies 1 and 2, the mind) than in a distal domain (in Studies 1 and 2, the brain). Study 3, therefore, asked participants to evaluate the effects of *two* real-life interventions, one acting on the mind (psychotherapy; as in Studies 1 and 2) and another acting on the brain (psychiatric medication).

Finally, in Study 3, we explored whether participants' prior exposure to academic psychology or neuroscience would predict the extent to which they endorse neurodualism versus the physicalist views typically embraced in the scientific literature. Exposure to scientific literature on the brain and the mind might be correlated with views more aligned with physicalism. However, given previous evidence that intuitive beliefs about minds and brains are not always affected by education (e.g., Ahn et al., 2017; Macdonald et al., 2017), we did not have strong a priori predictions in this respect. In addition, while education is an obvious source of information challenging popular dualist beliefs, it is not the only one: Claims and beliefs based on natural sciences and consistent with physicalism are ubiquitous in our culture. Even if a participant has not taken college-level classes in neuroscience or psychology, they may have received enough input elsewhere to suggest that the brain is not irrelevant for the mind.

4.1. Methods

4.1.1. Participants

Participants were recruited through MTurk, as in the preceding studies. Participants who had taken part in Studies 1 and 2 were prevented from participating. We recruited a total of 203 participants (75 females, 124 males, 1 other, 3 did not respond to demographic questions; mean age = 35.0 years; range = 19–70 years; 77.8% white; modal education = 4-year college degree). Of these, 39 participants were excluded for failing an attention check (same as in Study 1). We increased the sample size substantially from Study 2 to be able to investigate how individual differences in prior exposure to psychology and neuroscience relate to individual differences in participants' intuitive theories. A sensitivity power analysis (power = 80%, $\alpha = .05$, two-tailed tests) suggested that the sample in this study ($N = 164$ post-exclusions) was adequately powered to detect correlations as small as $r = .22$.

4.1.2. Materials and procedure

The materials and procedure were identical to Study 2 except for the following changes.

4.1.2.1. New task: Changes in the abstract. Participants' causal beliefs about mental and brain events in the abstract were assessed using a measure that consisted of two items: (1) "When there is a change in a person's mind, how often would you say there is also a change in the person's brain?" (mind-to-brain) and (2) "When there is a change in a person's brain, how often would you say there is also a change in the person's mind?" (brain-to-mind; see Table 1). Participants responded using a 0–100 sliding scale with the left and right poles labeled "NEVER" and "ALWAYS," respectively. The two items were presented in random order. We predicted a greater anticipated effect in the brain-to-mind direction than in the mind-to-brain direction.

4.1.2.2. Revisions to the locus-of-change task. In the revised locus-of-change task, the respondents were asked about two types of scenarios: ones in which the person saw a therapist (intervention on the mind), and ones in which they took medication (intervention on the brain; see Table 1). In every scenario, participants were also told that the person "completely overcame" their difficulties to underscore that there were no differences in treatment outcomes.[5] The scenarios and accompanying questions were as follows:

> Imagine the following scenario: A person has [difficulties in their personal life / been diagnosed with depression]. They [see a therapist / take medication] for [3 / 12] months and completely overcome their [problems / depression]. How much of a change would this treatment cause in the person's mind? And how much of a change would this treatment cause in the person's brain? Please click on both sliders to register your responses.

Of main interest to us was the interaction between the intervention (psychotherapy vs. medication) and the locus of change (proximal vs. distal). The neurodualism account predicts that participants will expect a smaller difference between proximal and distal effects for an intervention acting on the brain than for an intervention acting on the mind. Unlike in the changes-in-the-abstract and the hypothetical-alterations tasks, we do not make strong predictions about the distal effects of the two interventions because participants may have preconceived notions about the relative effectiveness of medications versus psychotherapy (Ahn et al., 2009, 2017; Angermeyer, van der Auwera, Carta, & Schomerus, 2017; Kemp et al., 2014; Khalsa et al., 2011), notions that may influence their responses even if they hear that the protagonist "completely overcame" their problem.

To avoid doubling the length of the task, we dropped the two of the problem types (professional difficulties and anxiety). In total, participants answered 2 (treatment: psychotherapy vs. medication) × 2 (effects: proximal vs. distal) × 2 (problems: personal difficulties vs. depression) × 2 (duration of treatment: 3 vs. 12 months) = 16 questions across 8 scenarios.

All psychotherapy scenarios were presented in one block, and all medication scenarios in another. The presentation order of the blocks and the items within each block was randomized across participants. The two treatment durations for a given scenario were presented in a set order with the 3-month item always presented first, before the 12-month item.

As in Study 2, the spatial arrangement of the response sliders was counterbalanced such that the mind-slider was presented above the brain-slider for half of the participants and below it for the other half. The order in which the mind and the brain were mentioned in the questions corresponded with the arrangement of the sliders, such that the question about the mind was always presented first whenever the mind-slider was presented first and vice versa.

The order of the new changes-in-the-abstract task and the locus-of-change task was randomized across participants, and the hypothetical-alterations task was always presented last.

4.1.2.3. Prior exposure. Participants were asked whether they had completed any college-level courses either in (a) psychology or (b) neuroscience and if so, how many courses (one vs. more than one).

4.2. Results

4.2.1. Changes-in-the-abstract task

Responses on this new task showed the predicted causal asymmetry: Participants believed that changes in the brain would more often be accompanied by changes in the mind (M = 67.6 [64.0, 71.2]) than vice versa (M = 56.2 [52.0, 60.3]), $F(1, 163) = 28.78$, $p < .001$, $\eta_p^2 = .150$.[6]

4.2.2. Locus-of-change task

Also consistent with neurodualism, participants reported that psychotherapy would exert much stronger proximal than distal effects (M = 75.4 [72.5, 78.3] vs. M = 47.0 [43.1, 50.8], respectively), whereas the anticipated proximal and distal effects of medication were virtually equal (M = 63.8 [60.0, 67.7] vs. M = 63.8 [60.1, 67.6]; see Figure 4). A 2 (intervention type)

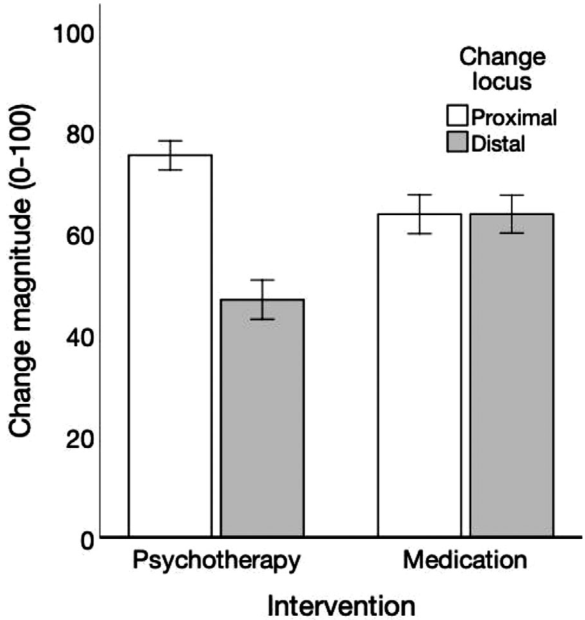

Fig 4. Participants' expectations of proximal and distal changes caused by psychotherapy (left) and psychiatric medication (right) in Study 3: Intervening on the brain affects the mind more than vice versa. Scale: 0 = "no change," 100 = "huge change." Error bars depict 95% confidence intervals.

×2 (locus of change) × 2 (problem) × 2 (treatment duration) repeated-measures ANOVA revealed the predicted two-way interaction between intervention type and locus of change, $F(1, 163) = 53.60$, $p < .001$, $\eta_p^2 = .247$.[7]

While not directly pertinent to our proposal, all four main effects were also significant in this ANOVA (*p*s < .026). Because the main effects of intervention type and locus of change are subsumed under the two-way interaction just described, we focus on the others: Participants expected interventions to have greater effects when they last 12 months ($M = 64.8$ [62.4, 67.2]) rather than just 3 months ($M = 60.2$ [57.8, 62.6]) and when treating depression ($M = 64.3$ [61.8, 66.8]) than problems in one's personal life ($M = 60.7$ [58.3, 63.1]). Finally, the three-way interaction among intervention, locus of change, and problem type was significant, $F(1, 163) = 6.52$, $p = .012$, $\eta_p^2 = .038$, suggesting that participants expected the distal versus proximal effects of psychotherapy and medications to be slightly different for depression and personal-life problems. This did not affect the predicted asymmetry or its interpretation, however.

4.2.3. Hypothetical-alterations task

As in previous studies, participants reported that an intervention by future scientists directly altering 25% of the brain would cause a substantially greater change in the mind ($M = 48.3$ [44.4, 52.1]) than altering the mind would cause in the brain ($M = 34.0$ [30.4, 37.5]), $F(1, 163) = 46.46$, $p < .001$, $\eta_p^2 = .222$.

4.2.4. Correlations across measures

Since the three tasks above are intended to tap the same intuitive theory, participants' responses to them should correlate. As shown in Table 1 in online Appendix B, participants' responses did in fact correlate across tasks quite reliably, and—as one would expect—the correlations were higher between questions that tapped intuitions about the same causal direction (e.g., the mind-to-brain questions on the locus-of-change and hypothetical-alterations tasks) than between questions that tapped intuitions about different causal directions (e.g., the mind-to-brain question on the locus-of-change task and the brain-to-mind question on the hypothetical-alterations task).

4.2.5. Prior exposure

Few participants had completed any college-level courses in neuroscience ($n = 12$ vs. 150 with no exposure), but many more had completed at least one psychology course ($n = 73$ vs. 89 with no exposure). Analyses of the three tasks showed that prior exposure to neither subject affected the observed asymmetry. Mixed ANOVAs, one with exposure to psychology and another with exposure to neuroscience as a dichotomous between-subjects variable, showed no interactions between any of the key measures and exposure to either subject on any of the three tasks (see Tables 2–7 in online Appendix B). The result regarding neuroscience, however, should be interpreted with caution, as few participants had taken neuroscience courses.

4.3. Discussion

Consistent with neurodualism, participants reported, across three different tasks, that events in the mind and brain have asymmetric consequences depending on the direction of causation. On the new task introduced in this study, participants reported that mental events are accompanied by brain events less often than vice versa. In and of itself, the fact that people believe mental events are accompanied by brain events only about half the time shows that people do not use a strictly physicalist model for reasoning about the mind and brain.

Study 3 also compared the anticipated effects of a psychological intervention to those of a biological one. Participants expected both interventions to have powerful proximal effects. Yet, while they believed the biological intervention would change the mind and brain equally, the anticipated effects of the psychological intervention were much weaker on the brain relative to its effects on the mind.

The present study also successfully replicated our previous findings regarding hypothetical alterations of the mind and the brain. Participants reported again that directly altering the brain would change the mind substantially more than vice versa.

Another noteworthy finding of this study was that participants' responses were systematically predictive of each other across tasks; these relationships were also sensitive to the direction of causal effects, meaning that correlations were higher among questions about the same causal direction than among questions about different causal directions. This pattern of results suggests that the three tasks provide valid measures of the hypothesized intuitive theory that people use when reasoning about minds and brains.

Finally, we found that prior exposure to neuroscience and psychology did not relate to the observed asymmetry in any of the three tasks, suggesting that the extent of exposure to scientific information (at least at the level assessed here) does not track with greater endorsement of physicalist views or lower endorsement of neurodualism. That exposure to college-level courses did not relate to the results is perhaps not surprising, considering that even academic neuroscientists do not consistently exhibit a fully physicalist world-view (Mudrik & Maoz, 2014). However, few of our participants had prior exposure to neuroscience; future research can hopefully address this question with greater statistical power.

5. Study 4: Revising the locus-of-change task to be more naturalistic; adding a measure of belief in popular dualism

A reasonable question about our studies so far concerns the ecological validity of our stimulus materials, as the scenarios described the problems to be treated only very briefly. Thus, in Study 4, we replaced the items in the locus-of-change task with more richly fleshed out character vignettes portraying named (fictional) individuals who are seeking help for specific mental health problems described in detail.

Study 4 also examined how participants' responses on the locus-of-change, hypothetical-alterations, and changes-in-the-abstract tasks relate to participants' endorsement of popular dualism (measured with the Dualism Scale devised by Stanovich, 1989; e.g., "The mind and the brain are two totally separate things"). To reiterate, the causal asymmetry observed in participants' responses in Studies 1–3 cannot have stemmed from popular dualism alone because this view does not license any systematic predictions about the relative causal powers of the brain and the mind. However, endorsement of popular dualism may still be correlated with responses on the three tasks, at least to an extent, since neurodualism is rooted partly in dualism.

5.1. Methods

5.1.1. Participants

Participants were recruited through MTurk, as in the preceding studies. Participants who had taken part in Studies 1–3 were prevented from participating. We recruited a total of 201 participants (109 females, 92 males; mean age = 35.5 years; range = 19–70 years; 70.6% white; modal education = 4-year college degree). Of these, 31 participants were excluded for failing an attention check (same as in Study 1).

5.1.2. Materials and procedure

The changes-in-the-abstract and hypothetical-alterations tasks were administered as in Study 3. Also as in Study 3, the changes-in-the-abstract and locus-of-change tasks were administered in randomized order, before the hypothetical-alterations task.

5.1.2.1. Revised locus-of-change task. As stimulus materials, we used four vignettes from Lebowitz and Ahn (2014), each describing an individual seeking help for a specific mental health problem (see Table 1). Two of the individuals exhibited symptoms fulfilling the criteria for depression and two for social phobia. Each participant viewed all four vignettes.

In each vignette, the symptoms were described in detail, but the diagnostic labels were not provided (for details, see Lebowitz & Ahn, 2014). One of the vignettes is shown below:

> Terry is a 28-year-old woman who is seeking treatment because she has felt deeply sad for the past 4 weeks. She reports that she has lost interest in activities that normally bring her pleasure, such as cooking and going to movies. She states that ever since the current period of sadness began, she has had difficulty with her memory and concentration. She has canceled plans with friends several times because she felt too tired to go out. She reports that her outlook on life has become quite pessimistic; she describes feeling hopeless about the future and perceiving herself to be totally worthless. She has reportedly experienced similar symptoms (a low mood that lasted several weeks) on three previous occasions in her life.

After the descriptions of the symptoms, the individual in the vignette was described as either seeing a therapist or taking medication, and the participants were told that after receiving this treatment for 12 months, their symptoms were now under control. After each vignette, the participants were asked two questions, using the "mental/brain processes" wording from Study 1B: (1) "How much of a change did this treatment cause in [the individual]'s mental processes?" and (2) "How much of a change did this treatment cause in [the individual]'s brain processes?"

Each participant was always presented with both treatment options for both depression and social phobia, in different vignettes. For instance, a participant told about one individual suffering from social phobia and receiving therapy in one vignette would always be presented with a vignette about another individual suffering from social phobia and receiving medication instead. The pairing of vignettes with treatments was counterbalanced across participants (e.g., 50% of participants read that a certain individual received therapy, and the others read that the same individual received medication). The presentation order of the vignettes was randomized for each participant. The response scales were the same as in Study 3 (i.e., separate 0–100 sliders for mind and brain, in counterbalanced order).

5.1.2.2. Dualism Scale. A subset of 10 items were selected from Stanovich's (1989) Dualism Scale. The items were chosen that had loadings 0.40 or higher on the factors labeled Classic Dualism and Rejection of Identity Theory, as reported by Stanovich (1989). Items included ones such as the following: "Minds are in principle independent of bodies, to which they are only temporarily 'attached'"; "My consciousness will survive the disintegration of my physical body"; and, "The mind and the brain are two totally separate things" (for the full list, see Table 3 in online Appendix C). Responses were elicited on a 1 ("strongly disagree") to 5 ("strongly agree") scale. Higher scores indicated higher endorsement of (popular)

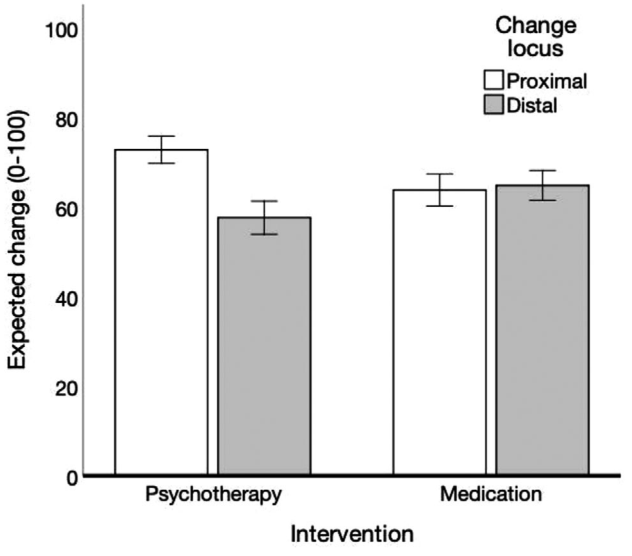

Fig 5. Participants' reports of anticipated psychotherapy- and medication-induced changes in Study 4. White and gray bars depict anticipated proximal and distal changes, respectively. Participants responded using a 0–100 slider with the left and right poles labeled "no change" and "huge change," respectively. Error bars depict 95% confidence intervals.

dualist beliefs. Responses were averaged across the 10 items (Cronbach's $\alpha = 0.70$). Due to an oversight, the scale was administered to only 162 participants (before exclusions).

5.2. Results

5.2.1. Changes-in-the-abstract task

As in Study 3 and consistent with neurodualism, participants reported that brain events would be accompanied by mental events ($M = 68.1$ [64.5, 71.7]) more often than vice versa ($M = 60.8$ [56.8, 64.9]), $F(1, 169) = 13.87$, $p < .001$, $\eta_p^2 = .076$.

5.2.2. Locus-of-change task

Despite the markedly different stimulus materials, the results were practically identical to those in Study 3 (see Figure 5). Participants again reported that a psychological intervention (psychotherapy) would exert much stronger proximal than distal effects ($M = 72.8$ [69.8, 75.8] vs. $M = 57.6$ [53.9, 61.3], respectively), whereas a biological intervention acting on the brain (medication) would produce more equal proximal and distal effects ($M = 63.8$ [60.2, 67.4] vs. $M = 64.8$ [61.5, 68.2], respectively). In a 2 (intervention type) $\times$ 2 (locus of change) $\times$ 2 (problem) repeated-measures ANOVA, the predicted interaction between intervention type (psychotherapy vs. medication) and locus of change (proximal vs. distal) was significant, $F(1, 168) = 23.64$, $p < .001$, $\eta_p^2 = .123$.

We also found a three-way interaction among intervention type, locus of change, and problem type, $F(1, 168) = 5.05$, $p = .026$, $\eta_p^2 = .029$, reflecting slight differences in beliefs about effects for depression versus social phobia.[8] These subtle differences, however, did not affect the asymmetry that was our main interest. No other effects were significant (apart from a main effect of locus of change, qualified by the two-way interaction).

5.2.3. Hypothetical-alterations task

As in Study 3 and consistent with neurodualism, participants reported that directly altering 25% of the brain would cause a greater change in the mind ($M = 51.4$ [47.8, 55.1]) than altering 25% of the mind would cause in the brain ($M = 46.1$ [42.1, 50.1]), $F(1, 166) = 9.23$, $p = .003$, $\eta_p^2 = .053$.

5.2.4. Relation to endorsement of popular dualism

The Dualism Scale was only weakly predictive of responses on the three tasks, with .17 being the highest and the only significant correlation (see Tables 1 and 2 in online Appendix C). In an additional set of analyses, we divided participants into a high-dualism and low-dualism group using a median split (for descriptive statistics, see online Appendix C) and added this dualism factor to the ANOVAs used to analyze responses on each task. The dualism variable (high vs. low) did not moderate any of the key main effects or interactions (i.e., those indicating a causal asymmetry) across the three tasks (ps $> .45$).

5.3. Discussion

Study 4 successfully replicated the findings from the locus-of-change task from Study 3 despite considerably different stimulus materials. Participants again believed, when presented with detailed character vignettes describing people seeking help for specific mental health problems, that an intervention on the brain (psychiatric medications) would produce more similar proximal and distal effects than an intervention on the mind (psychotherapy) would. These results rule out the possibility that the pattern in the previous studies was limited to a particular type of materials or level of abstraction in the stimuli and suggest that the observed pattern is fairly robust. Notably, the results on the changes-in-the-abstract and hypothetical-alterations tasks also replicated those of Studies 1–3.

Finally, participants' endorsement of simple dualist beliefs, as measured by Stanovich's (1989) Dualism Scale, was not predictive of nearly any of the responses on our tasks. This result supports our premise that neurodualism does not boil down to popular dualism, but comprises a more complex set of beliefs.

6. Study 5 (preregistered): Do psychotherapists endorse neurodualism?

Do trained psychotherapists also intuitively believe, like lay participants in Studies 1–4, that the brain is more powerful in changing the mind than the mind is in changing the brain? Although mental health professionals are sometimes prone to the same biases as lay participants (e.g., Ahn et al., 2009, 2017), psychotherapists provide a conservative test of neurodualism: As mental health professionals, psychotherapists are likely to be better informed about the scientific evidence regarding, for example, psychotherapy-related changes in the brain (e.g., Linden, 2006), which could promote a wider appreciation of the distal effects of psychological interventions. In addition, because they themselves are providers of psychotherapy, psychotherapists are conceivably more likely than lay participants to believe that psychotherapy has powerful causal effects, perhaps including distal ones on the brain (contrary to neurodualism).

We also explored whether neurodualism among psychotherapists is associated with particular patterns of treatment recommendations. If providers of psychological treatments themselves regard the brain as more causally powerful, this intuition could inflate the anticipated effectiveness of biological treatments for mental health conditions, which could in turn affect clinical judgment and treatment recommendations.

6.1. Methods

6.1.1. Participants

The participants were randomly selected from a national database (the American Psychological Association's licensed psychologist locator, https://locator.apa.org/). We randomized the order of all ZIP codes in the United States and contacted clinicians working within a 25-mile radius from each ZIP code on this randomized list until we reached the desired sample size. No more than 20 clinicians were selected from any single ZIP code, even if the locator listed more. Because the 25-mile radius yielded a substantial number of duplicates, we decreased the radius to 10 miles toward the end of the recruitment process. Participants were not paid for responding. Of the 998 psychotherapists contacted, a total of 117 participants (11.7%) met our preregistered inclusion criterion (i.e., they responded to all questions in the locus-of-change task).

To arrive at the preregistered sample size of 170 (identical to that in Study 4), we conducted a second wave of data collection through the same psychologist locator, this time offering $10 Amazon.com gift cards in exchange for participation. The experiment was closed when the total sample exceeded the targeted count in the preregistered plan. A total of 514 psychotherapists were contacted in this second wave, and 63 of them participated (12.3%). Two respondents were excluded for having practiced for 0 years, as specified in our preregistered protocol,[9] bringing the total number of participants (post-exclusions) to 178.

Most respondents had a PhD (66.9%) or a PsyD (26.4%). Nine (5.1%) had an EdD, and two (1.1%) an MSW. In addition, 11 (6.2%) had also completed another degree, such as MA, MS, or an MBA. None of the participants had an MD, but one respondent (0.6%) reported having the right to prescribe psychiatric medications[10] (six did not respond). The psychotherapists practiced in 29 states, with California (20.8% of respondents), New York (10.7%), and Illinois (7.9%) the most prominently represented. They had practiced for a median of 21.5 years (mode = 15, range = 0.75–50). As their main approaches or orientations with their clients, they reported behavioral/cognitive-behavioral (49.4%), psychoanalysis/psychodynamic (31.5%), integrative/holistic (28.7%), interpersonal (28.1%), family systems (18.0%), cognitive (18.0%), humanistic/Gestalt/existential (14.6%), and other (17.4%). (Participants could select multiple options on this item.)

The respondents ranged in age from 28 to 85 years (median = 58.0), and their gender identities were 57.9% women, 37.6% men, and 1.1% other (3.4% did not respond). The participants were predominantly white (88.2%), with 3.9% Asian/Asian American, 2.2% Hispanic/Latinx, 0.6% Black/African American, and 3.4% more than one race (3.9% did not respond).

6.1.2. Materials and procedure

The locus-of-change, hypothetical-alterations, and changes-in-the-abstract tasks were administered exactly as in Study 4 (see Table 1). To investigate whether psychotherapists' intuitive theories were associated with real-life clinical decisions, at the end of the sessions, we asked respondents to indicate what percentage of clients diagnosed with (a) depression and (b) anxiety they referred for prescriptions (or prescribed medication to, where applicable).

6.1.3. Preregistration

The analysis plan and hypotheses were preregistered on AsPredicted.org: https://aspredicted.org/vs9m8.pdf

6.2. Results

The results resembled those with lay participants, but in an attenuated form. In two tasks out of three tasks (locus-of-change and hypothetical-alterations), the psychotherapists showed a similar causal asymmetry as untrained participants. In contrast, the results from the changes-in-the-abstract task differed from the results with lay participants.

6.2.1. Locus-of-change task

Not surprisingly for providers of psychotherapy, the participants reported that psychotherapy (M = 63.5 [61.0, 66.0]) would cause greater changes overall than medications (M = 51.7 [48.8, 54.7]), $F(1, 177) = 74.25$, $p < .001$, $\eta_p^2 = .296$. In addition, the psychotherapists believed that, overall, treatments would cause greater changes proximally (M = 66.1 [63.6, 68.6]) than distally (M = 49.1 [46.2, 52.1]), $F(1, 177) = 158.28$, $p < .001$, $\eta_p^2 = .472$. Importantly, however, these main effects were qualified by a significant interaction between intervention type and locus of change, consistent with neurodualism (see Figure 6): Psychotherapy providers also believed, like lay participants, that the difference between proximal and distal causal effects is smaller for medication relative to psychotherapy, $F(1, 177) = 10.34$, $p = .002$, $\eta_p^2 = .055$.[11]

6.2.2. Hypothetical-alterations task

Like lay participants, psychotherapists expected that changing the brain would have a stronger effect on the mind (M = 43.7 [39.7, 47.7]) than vice versa (M = 40.5 [36.8, 44.3]). Unlike with lay participants, however, this difference was only marginally significant, $F(1, 172) = 3.36$, $p = .069$, $\eta_p^2 = .019$. This suggests that if psychotherapists do exhibit the same asymmetry as lay participants, its magnitude may be weaker.

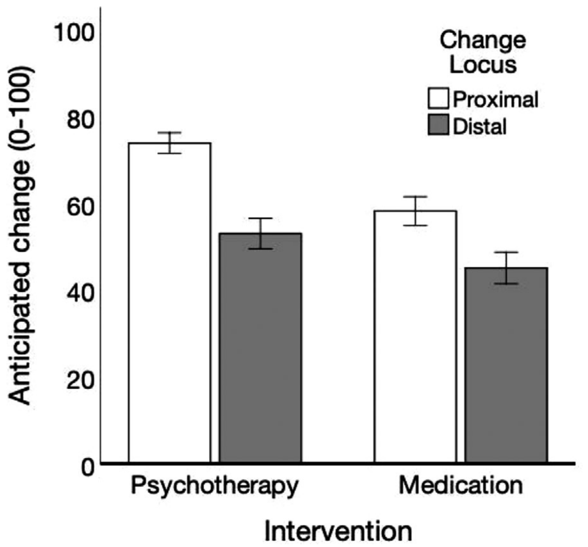

Fig 6. Psychotherapists' reports of anticipated proximal and distal effects of psychotherapy (left) and medication (right) in Study 5. Professional providers of psychotherapy expected both psychotherapy and medication to exert greater changes in their proximal (white bars) than distal (gray bars) metaphysical domain. However, the difference between proximal and distal effects was smaller for medication than for psychotherapy, suggesting a causal asymmetry in intuitive theories of the mind and the brain (neurodualism). Error bars depict 95% confidence intervals.

6.2.3. Changes-in-the-abstract task

The pattern of results was clearly discrepant from lay participants in the changes-in-the-abstract task. Contrary to our preregistered hypothesis, the psychotherapists expected changes in the mind to be accompanied by changes in the brain *more often* than vice versa ($M = 67.3$ [63.6, 70.9] vs. $M = 60.4$ [56.8, 64.0], respectively), $F(1, 176) = 12.57, p < .001, \eta_p^2 = .067$. The psychotherapists, however, like lay participants, also diverged from a strictly physicalist position, in that they did not believe that changes in the mind were always accompanied by changes in the brain (i.e., 67.3 is substantially below 100), consistent with prior findings suggesting that trained mental health clinicians also endorse dualistic beliefs (Miresco & Kirmayer, 2006).

6.2.4. Relation to treatment preferences

Next, we tested whether psychotherapists' intuitive theories of the relation between the mind and the brain predicted their treatment preferences—specifically, how often they referred patients diagnosed with depression or anxiety for medication prescriptions or prescribed medications themselves to such patients (if they had the right to do so).

We first calculated an index of neurodualism from each of the three tasks by subtracting the anticipated changes to the brain induced by interventions on the mind from the anticipated changes to the mind of an analogous intervention on the brain. Higher values on these three indices corresponded to a greater endorsement of neurodualism. Because these three indices were correlated, rs $> .41$, ps $< .001$, suggesting that they all tapped the same intuitive theory, we averaged them into a composite index of neurodualism (Cronbach's $\alpha = .71$).

Similarly, because the percentages of patients with depression ($M = 43.6\%$, mode $= 50\%$) and anxiety ($M = 32.9\%$, mode $= 50\%$) that the psychotherapists referred for prescriptions were also correlated, $r(170) = .75$, $p < .001$, we averaged them into a composite index of preference for psychiatric medications.

Finally, we examined the correlation between the psychotherapists' endorsement of neurodualism and their preference for medications. This correlation was significant, $r(170) = .19$, $p = .015$. Psychotherapists who believed the brain affects the mind more strongly than the mind affects the brain were also more likely to encourage their patients to take psychiatric drugs. This relationship remained significant when simultaneously adjusting for the psychotherapists' years of experience, age, and gender, $\beta = .22$, $p = .004$.

6.3. Discussion

In Study 5, we investigated whether professional psychotherapists operate with the same intuitive theory of the mind and the brain that we found among lay participants in Studies 1–4. Indeed, in two out of three tasks, trained psychotherapy providers showed some evidence of neurodualism, but to a weaker extent than lay participants. Arguably, these results highlight how intuitive and pervasive the neurodualist way of thinking about the mind and the brain is.

Perhaps the clearest evidence of neurodualism was observed in the locus-of-change task, where—like lay participants—psychotherapists expected the difference between proximal and distal causal effects to be smaller for medication relative to psychotherapy. Even on this task, however, psychotherapists' responses were subtly different from those of lay participants. The most noticeable difference concerned the anticipated distal effects of medications on the mind (compare the rightmost bars of Figures 5 and 6): Psychotherapists did not believe that medications would affect the mind to the same extent as lay participants believed they would. This is perhaps not surprising from providers of a competing form of treatment. In fact, an exploratory post-hoc analysis suggested that it was precisely the therapists who were leery of medications (as evidenced by their referral history) who were also skeptical of the effects of medications on the mind (see online Appendix D.2). Several participants even explicitly articulated a belief that drugs do not cause lasting psychological change: for example, "While medication may have an impact during its usage, it does not make permanent changes—that is a different question—whereas I believe therapy and experience DO have the ability to cause permanent changes."

Perhaps the same reasons also explain why psychotherapists showed the opposite pattern to that of lay participants on the changes-in-the-abstract task, expecting mental events to be accompanied by neural events more often than vice versa: Psychotherapists may be particularly inclined to believe that psychological interventions can produce profound and lasting effects, and that biological interventions may not always produce wanted mental changes. This belief may have guided their reasoning in this thought experiment as well. Indeed, an exploratory analysis reported in online Appendix D.2 suggests that psychotherapists who were wary of psychiatric drugs exhibited the pattern of responses that was most discrepant from that observed among lay participants. We should also note that it is a scientific fact that many neural events simply have to do with the functioning of the body (e.g., the neural activity that regulates our breathing while we sleep) and thus lack an intuitively recognizable mental counterpart, whereas most scientists would agree that any mental event must have a corresponding neural event. Perhaps recognition of this fact also played a part in psychotherapists' responses on the changes-in-the-abstract task.

Our other goal in Study 5 was to investigate whether psychotherapists' intuitive theories of the relation between the mind and the brain predict a key aspect of how they practice their profession—that is, the extent to which they recommend to their patients that they take psychiatric medications. We found evidence of such a relationship: Responses on the neurodualism tasks predicted treatment preferences, even when adjusting for covariates such as years of experience. This result suggests that consequential aspects of how mental health professionals work with their patients are informed by an intuitive neurodualist theory of mind–brain interactions that is out of step with current scientific understanding.

7. General discussion

We investigated intuitive theories of minds and brains in five studies with both lay participants and professional psychotherapists. We hypothesized that when reasoning about minds and brains, people rely on neurodualism—a hybrid intuitive theory that assimilates aspects of physicalist beliefs into pre-existing dualist intuitions, attributing more causal power to the brain over the mind than vice versa.

In all experiments and across several different tasks involving both thought experiments and naturalistic scenarios, untrained participants believed that interventions acting on the brain would affect the mind more than interventions acting on the mind would affect the brain, supporting our proposal. This causal asymmetry was strong and replicated reliably with untrained participants. Moreover, the extent to which participants endorsed popular dualism was only weakly correlated with their endorsement of neurodualism, supporting our proposal that a more complex set of beliefs is involved. In the last study, professional psychotherapists also showed evidence of endorsing neurodualism—albeit to a weaker degree—despite their scientific training and stronger reluctance, relative to lay participants, to believe that psychiatric medications affect the mind.

Our results both corroborate and extend prior findings regarding intuitive reasoning about minds and brains. Our results *corroborate* prior findings by showing, once again, that both laypeople and trained mental health professionals commonly hold dualistic beliefs. If their reasoning had been based on (folk versions of) a physicalist model, such as identity theory or supervenience, participants should not have expected mental events to occur in the absence of neural events. However, both lay participants and professional psychotherapists did consistently report that mental changes can occur (at least sometimes) even in situations in which no neural changes occur.

Our findings also *extend* prior findings by demonstrating that intuitive theories of minds and brains are considerably more complex than has previously been acknowledged. While it is widely agreed that dualistic beliefs are common (Ahn et al., 2017; Bloom, 2004; Forstmann & Burgmer, 2015; Miresco & Kirmayer, 2006; Mudrik & Maoz, 2014; Stanovich, 1989), how *exactly* people reason about the mind and brain in relation to each other has remained unclear. Our findings show that the fuller picture of intuitive theories is more nuanced than a mere belief that the mind and the brain are separate interacting entities. That intuitive theories can contain aspects of both popular-dualist and physicalist beliefs helps to explain why people's beliefs often seem internally inconsistent: While people often agree with the statement that the mind is not separable from the brain, they also endorse the view that the mind is *not* fundamentally physical (Demertzi et al., 2009). Similarly, even professional neuroscientists—who presumably endorse physicalist views—commonly discuss the brain in terms that conflict with physicalism (Greene, 2011; Mudrik & Maoz, 2014). Inconsistencies such as these are to be expected if people intuitively think of the mind as neither purely physical nor entirely independent of the brain, but rather embrace aspects of both of those views simultaneously. In fact, it is not uncommon for intuitive theories to take the form of hybrids that incorporate novel beliefs into existing theories whose original core is not lost even as the theories become increasingly complex (e.g., Hussak & Cimpian, 2019; Shtulman & Lombrozo, 2016).

Moreover, the current study sheds light on what this hybrid theory looks like. The results suggest that even if (and when) people are dualists, they perceive the brain neither as causally irrelevant for the mind nor as unresponsive to mental changes, but rather see the brain as a more commanding and robust causal agent than the mind. Future research will hopefully be able to capture further subtleties in intuitive theories of minds and brains. It seems likely that if researchers search for more fine-grained options than dichotomous dualist/antidualist positions in lay intuitions, increasingly fine-grained aspects may become visible.

7.1. Broader implications for theory and practice

7.1.1. Relation to the popular allure of neuroscience

Our findings may help to explain the intense fascination that the general public and mass media show for neuroscience research (Beck, 2010; O'Connor, Rees, & Joffe, 2012). If the general public is reluctant to believe that changes in the mind always correspond to changes in the brain, neuroscience findings showing that what happens in our minds happens in our brains as well contradict this belief and may thereby be particularly intriguing. Neurodualism may also help explain why people find brain-related statements informative in the context of psychological explanations even when the statements are irrelevant (Weisberg, Keil, Goodstein, Rawson, & Gray, 2008; Fernandez-Duque, Evans, Christian, & Hodges, 2015). Conceivably, the intuitive tendency to privilege causal patterns in the brain-to-mind direction (i.e., neurodualism) may bias people to perceive causal brain-to-mind connections even when none exist, which may in turn make the addition of neuroscience evidence to a psychological explanation *seem* informative. Also consistent with this argument, Fernandez-Duque et al. (2015) found that (popular) dualistic beliefs alone did not predict their participants' reasoning in these contexts. In future research, it would be useful to test whether endorsement of neurodualism *does* predict the tendency to view information about the brain as particularly explanatory even in cases where it is not.

On a different note, some authors have suggested that the allure of neuroscience explanations is not specific to beliefs about minds and brains but related to a more general preference for reductive information. Hopkins, Weisberg, and Taylor (2016) found that across different scientific disciplines, people generally preferred explanations that referred to processes perceived as more fundamental, even when these processes were logically irrelevant to the explanation. According to this view, information about the brain may be seen as particularly informative because it is perceived as operating at the next level of analysis below psychological phenomena (Fernandez-Duque, 2017). It seems likely, however, that the neurodualist intuitive theory identified in the present research and this general preference for reductive information are independent inputs into the public's fascination with neuroscience explanations. Importantly, neurodualism itself is not a reductionist theory: For instance, people report that changes in mental states are only *sometimes* accompanied by changes in brain states (see Studies 3–5). Beliefs such as these are not easily interpreted as evidence that people are treating the terms "mind" and "brain" as referring to the same phenomenon at different levels of analysis. A more plausible account is that a neurodualist intuitive theory and the preference for reductive explanations are two *independent* factors contributing to the public appeal of neuroscience.

7.1.2. Implications for reasoning about mental illness and health

The current results may help to make sense of common beliefs regarding treatment efficacy in mental health. When people think that the source of a mental health issue such as depression is in the brain, they perceive psychological interventions as less likely to be helpful (Ahn et al., 2017; Deacon & Baird, 2009; Kemp et al., 2014). The belief that a psychological treatment cannot be effective if the problem is reflected in brain processes is at odds with both a physicalist view of the mind and the empirical evidence (e.g., Linden, 2006; Lozano, 2011; Deacon, 2013). These beliefs are unfortunate from a practical viewpoint as well because prognostic beliefs often predict treatment outcomes (Rutherford, Wager, & Roose, 2010). That is, pessimistic expectancies can become self-fulfilling prophecies: Neurobiological causal attributions are associated with both lower treatment expectations and poorer psychosocial treatment outcomes in depression (Schroder et al., 2020). Our findings suggest that part of the reason for these effects may lie in the intuitive theories people use for reasoning about the mind and brain. Biological causal explanations may foster pessimism about the efficacy of psychotherapy partly because of an underlying intuitive theory ascribing relatively little power to the mind over the brain.

Fortunately, targeted education about the malleability of neurobiological factors in depression can help reduce prognostic pessimism and strengthen patients' beliefs about their own ability to regulate their moods in depression (Lebowitz & Ahn, 2015), suggesting that these intuitions are not fixed or immutable. In future work, it would be worthwhile to investigate whether interventions that target people's intuitive theories of the relation between the mind and brain could also help mitigate the negative consequences of biological attributions for disorders such as depression.

While participants in our studies were reluctant to believe that acting on the mind can result in changes in the brain, they were more willing to endorse that acting on the brain can result in changes in the mind. This may help, in part, to explain why Western societies have so enthusiastically come to favor neurobiologically centered approaches to mental illness despite people's dualistic intuitions. Pharmacological treatments have become the predominant societal response to mental health conditions over the past decades. Although it is widely agreed that an adequate response to mental distress needs to address several nonreducible levels, Western cultures have allowed "the biopsychosocial model to become the bio-bio-bio model," in the words of a previous president of the American Psychiatric Association (Sharfstein, 2005). Arguably, neither the enthusiasm nor the scale at which this approach has been implemented is easy to explain from a purely evidence-based perspective (Deacon, 2013; Whitaker & Cosgrove, 2015; UN Human Rights Council, 2017; Lacasse & Leo, 2005; Healy, 2015; Moncrieff & Cohen, 2006), and its success has been controversial at best (Danborg & Gøtzsche, 2019; Gøtzsche, Young, & Crace, 2015; Haslam & Kvaale, 2015; Hengartner, 2020; Ioannidis, 2019; Jakobsen et al., 2017; Munkholm, Paludan-Müller, & Boesen, 2019; Sohler et al., 2015). Why, then, do we continue to operate based "on faith that neuroscience will eventually revolutionize mental health practice," if "[d]ecades of extraordinary investment in biomedical research have not been rewarded with improved clinical tools or outcomes" (Deacon, 2013, p. 858)? While numerous societal and institutional factors undoubtedly affect the situation in all its complexity (e.g., Moncrieff, 2006; Whitaker & Cosgrove, 2015), from a strictly cognitive perspective, it is conceivable that our intuitive theories—in particular, our willingness to believe in the brain as an asymmetrically powerful causal agent that can influence the mind—may have contributed and made the public prone to believe overstated neuroscientific claims. In a self-reinforcing cycle, the widescale implementation of any neurobiologically centered practices likely also loops back and shapes people's intuitive theories in ways that further increase the appeal of these practices.

7.1.3. Relation to the broader historical context

The intuitive theories documented here are undoubtedly a product of the current historical context: Several authors have suggested that many cultures are undergoing a transition toward understanding mind–brain relations in more materialistic terms (e.g., Mudrik & Maoz, 2014). As scientific inquiry has progressed, we as a culture have increasingly come to believe that it is the brain which controls faculties formerly associated with the soul, such as memory, language, and emotion. If the suggestion is correct that we are in the process of intuitively giving up the mind's and/or soul's functions to material brains (Greene, 2011), it is interesting to consider what are "the soul's last stands"—the most immaterial of our nonphysical capacities, the ones not yet outsourced to the brain.

7.2. Conclusion

It is important to keep in mind that, philosophically, the mind–body problem remains an unresolved paradox. Although materialist and physicalist views have been the working assumption of contemporary psychologists and neuroscientists and also the prevailing position in philosophy over the past decades, this does not mean that the original mind–body problem itself was resolved. It remains, to this day, extremely difficult to see how, if the mind is a nonphysical thing and the body is a physical thing, one could simply just be the other (or how they could interact, if we are dualists). It is helpful to remember that not only the general public but also (at least some) contemporary philosophers find the claim inherently implausible that the mind simply *is* a physical thing (e.g., Westphal, 2016). What people *think* the mind is, however, and how exactly they think it is related to the brain seems worth investigating further, both for theoretical and practical reasons.

Notes

1 Although materialism and physicalism differ in their nuances (Stoljar, 2017), we will use the terms interchangeably in this manuscript.

2 In this context, psychiatric medications can be thought of as acting *primarily* on the brain. Taking drugs of course *also* has powerful additional psychological effects, as attested by extensive research literatures on placebo and nocebo effects. Precisely because such complications are unavoidable in naturalistic scenarios, it is important to complement them with conceptually tidier thought experiments.

3 Unless mental events are thought not to exist at all, or to be misleading about actual physical events in the brain (see Churchland, 1981; Dennett, 1989). Most philosophers and neuroscientists, however, take less radical positions regarding commonsense views of mental phenomena.

4 The Huynh–Feldt ($\varepsilon = .83$) correction was used because the assumption of sphericity was violated, $\chi^2(5) = 16.92$, $p < .01$.

5 It is important to stipulate identical outcomes for the two treatments in case participants have prior beliefs about the efficacy of psychotherapy versus psychiatric medication, which would be reflected in their change estimates.

6 To investigate the possibility of carryover effects from the locus-of-change task, we analyzed separately the responses from the subset of participants for whom the changes-in-the-abstract task was presented first ($n = 81$). The responses of this subset of participants showed the same asymmetry, $M = 67.9$ [62.4, 73.4] versus 59.9 [53.6, 66.2], $F(1, 80) = 6.00$, $p = .016$, $\eta_p^2 = .070$. Thus, the predicted asymmetry on this task was not an artifact of task order.

7 To investigate the possibility of carryover effects from the change-in-the abstract task, we analyzed separately the responses from the subset of participants for whom the locus-of-change task was presented first ($n = 83$). The responses of this subset of participants showed the same predicted two-way interaction between intervention type and locus of change, $F(1, 82) = 45.93$, $p < .001$, $\eta_p^2 = .359$. Thus, the predicted asymmetry on this task was not an artifact of task order.

8 For social phobia, the anticipated proximal versus distal effects were $M = 73.9$ [70.7, 77.0] versus $M = 57.5$ [53.5, 61.4] for therapy and $M = 63.4$ [59.5, 67.3] versus $M = 66.6$ [63.1, 70.2] for medication. For depression, the anticipated proximal versus distal effects were $M = 71.7$ [68.2, 75.3] versus $M = 57.8$ [53.7, 61.9] for therapy and $M = 64.2$ [60.4 vs. 68.1] versus $M = 63.1$ [59.3, 66.8] for medication.

9 Four other participants did not respond to the question asking how long they had practiced. They were included in all analyses.
10 Clinical psychologists have the right to prescribe psychiatric medications in some states.
11 In addition, there was a significant interaction between intervention and problem type (anxiety vs. depression). As this was not directly relevant for our hypotheses, the details are described in online Appendix D.1.

References

Ahn, W. (1998). Why are different features central for natural kinds and artifacts? The role of causal status in determining feature centrality. *Cognition*, *69*, 135–178.

Ahn, W., Kim, N. S., & Lebowitz, M. S. (2017). The role of causal knowledge in reasoning about mental disorders. In M. R. Waldmann (Ed.), *The Oxford handbook of causal reasoning* (pp. 603–617). https://doi.org/10.1093/oxfordhb/9780199399550.013.31

Ahn, W., Proctor, C. C., & Flanagan, E. H. (2009). Mental health clinicians' beliefs about the biological, psychological, and environmental bases of mental disorders. *Cognitive Science*, *33*(2), 147–182.

Angermeyer, M. C., van der Auwera, S., Carta, M. G., & Schomerus, G. (2017). Public attitudes towards psychiatry and psychiatric treatment at the beginning of the 21st century: A systematic review and meta-analysis of population surveys. *World Psychiatry*, *16*(1), 50–61.

Beck, D. M. (2010). The appeal of the brain in the popular press. *Perspectives on Psychological Science*, *5*(6), 762–766.

Bering, J. M. (2006). The folk psychology of souls. *Behavioral and Brain Sciences*, *29*, 453–498.

Bloom, P. (2004). *Descartes' baby: How the science of child development explains what makes us human*. New York: Basic Books.

Bloom, P. (2007). Religion is natural. *Developmental Science*, *10*(1), 147–151. https://doi.org/10.1111/j.1467-7687.2007.00577.x

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Chudek, M., McNamara, R. A., Birch, S., Bloom, P. & Henrich, J. (2018). Do minds switch bodies? Dualist interpretations across ages and societies. *Religion, Brain & Behavior*, *8*(4), 354–368. https://doi.org/10.1080/2153599X.2017.1377757

Churchland, P. M. (1981). Eliminative materialism and propositional attitudes. *Journal of Philosophy*, *78*(2), 67–90.

Churchland, P. M. (2013). Matter and consciousness *(3rd ed.)*. Cambridge, MA: MIT Press.

Danborg, P. B., & Gøtzsche, P. C. (2019). Benefits and harms of antipsychotic drugs in drug-naïve patients with psychosis: A systematic review. *International Journal of Risk & Safety in Medicine*, *30*, 193–201. https://doi.org/10.3233/JRS-195063

Deacon, B. J. (2013). The biomedical model of mental disorder: A critical analysis of its validity, utility, and effects on psychotherapy research. *Clinical Psychology Review*, *33*(7), 846–861.

Deacon, B. J., & Baird, G. L. (2009). The chemical imbalance explanation of depression: Reducing blame at what cost? *Journal of Social and Clinical Psychology*, *28*(4), 415–435.

Demertzi, A., Liew, C., Ledoux, D., Bruno, M. A., Sharpe, M., Laureys, S., & Zeman, A. (2009). Dualism persists in the science of mind. *Annals of the New York Academy of Sciences*, *1157*, 1–9. https://doi.org/10.1111/j.1749-6632.2008.04117.x

Dennett, D. C. (1989). *The intentional stance*. Cambridge, MA: MIT Press.

Fernandez-Duque, D. (2017). Lay theories of the mind/brain relationship and the allure of neuroscience. In C. M. Zedelius, B. Müller, & J. W. Schooler (Eds.), *The science of lay theories: How beliefs shape our cognition, behavior, and health* (pp. 207–227). Cham: Springer. https://doi.org/10.1007/978-3-319-57306-9_9

Fernandez-Duque, D., Evans, J., Christian, C., & Hodges, S. D. (2015). Superfluous neuroscience information makes explanations of psychological phenomena more appealing. *Journal of Cognitive Neuroscience*, *27*(5), 926–944.

Forstmann, M., & Burgmer, P. (2015). Adults are intuitive mind-body dualists. *Journal of Experimental Psychology: General*, *144*(1), 222–235. https://doi.org/10.1037/xge0000045

Forstmann, M., Burgmer, P., & Mussweiler, T. (2012). "The mind is willing, but the flesh is weak": The effects of mind-body dualism on health behavior. *Psychological Science*, *23*(10), 1239–1245.

Gerstenberg, T., & Tenenbaum, J. B. (2017). Intuitive theories. In M. R. Waldmann (Ed.), *Oxford handbook of causal reasoning* (pp. 515–548). New York: Oxford University Press.

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind & Language*, *7*, 145–171. https://doi.org/10.1111/j.1468-0017.1992.tb.00202.x

Gøtzsche, P. C., Young, A. H., & Crace, J. (2015). Does long term use of psychiatric drugs cause more harm than good? *BMJ*, *350*, h2435. https://doi.org/10.1136/bmj.h2435

Greeley, A. M., & Hout, M. (1999). Americans' increasing belief in life after death: Religious competition and acculturation. *American Sociological Review*, *64*(6), 813–835. https://doi.org/10.2307/2657404

Greene, J. D. (2011). Social neuroscience and the soul's last stand. In A. Todorov, S. Fiske, & D. Prentice (Eds.), *Social neuroscience: Toward understanding the underpinnings of the social mind* (pp. 263–273). New York: Oxford University Press.

Haslam, N., & Kvaale, E. P. (2015). Biogenetic explanations of mental disorder: The mixed-blessings model. *Current Directions in Psychological Science*, *24*(5), 399–404. https://doi.org/10.1177/0963721415588082

Healy, D. (2015). Serotonin and depression: The marketing of a myth. *BMJ*, *350*, h1771. https://doi.org/10.1136/bmj.h1771

Hengartner, M. P. (2020). How effective are antidepressants for depression over the long term? A critical review of relapse prevention trials and the issue of withdrawal confounding. *Therapeutic Advances in Psychopharmacology*, *10*, 1–10. https://doi.org/10.1177/2045125320921694

Hodge, K. M. (2008). Descartes' mistake: How afterlife beliefs challenge the assumption that humans are intuitive Cartesian substance dualists. *Journal of Cognition and Culture*, *8*, 387–415.

Hopkins, E. J., Weisberg, D. S., & Taylor, J. C. V. (2016). The seductive allure is a reductive allure: People prefer scientific explanations that contain logically irrelevant reductive information. *Cognition*, *155*, 67–76. https://doi.org/10.1016/j.cognition.2016.06.011

Horne, Z., & Cimpian, A. (2019). Intuitions about personal identity are rooted in essentialist thinking across development. *Cognition*, *191*, 103981.

Hussak, L. J., & Cimpian, A. (2019). "It feels like it's in your body": How children in the United States think about nationality. *Journal of Experimental Psychology: General*, *148*(7), 1153–1168.

Inagaki, K., & Hatano, G. (1993). Young children's understanding of the mind-body distinction. *Child Development*, *64*(5), 1534–1549.

Ioannidis, J. P. A. (2019). Therapy and prevention for mental health: What if mental diseases are mostly not brain disorders? *Behavioral and Brain Sciences*, *42*, e13. https://doi.org/10.1017/S0140525X1800105X

Jakobsen, J. C., Katakam, K. K., Schou, A., Hellmuth, S. G., Stallknecht, S. E., Leth-Møller, K., Iversen, M., Banke, M. B., Petersen, I. J., Klingenberg, S. L., Krogh, J., Ebert, S. E., Timm, A., Lindschou, J., & Gluud, C. (2017). Selective serotonin reuptake inhibitors versus placebo in patients with major depressive disorder. A systematic review with meta-analysis and trial sequential analysis. *BMC Psychiatry*, *17*(1), 58.

Johnson, C. N., & Wellman, H. M. (1982). Children's developing conceptions of the mind and brain. *Child Development*, *53*(1), 222–234.

Kemp, J. J., Lickel, J. J., & Deacon, B. J. (2014). Effects of a chemical imbalance causal explanation on individuals' perceptions of their depressive symptoms. *Behaviour Research and Therapy*, *56*, 47–52.

Khalsa, S. R., McCarthy, K. S., Sharpless, B. A., Barrett, M. S., & Barber, J. P. (2011). Beliefs about the causes of depression and treatment preferences. *Journal of Clinical Psychology*, *67*(6), 539–549.

Kim, J. (1998). *Mind in a physical world: An essay on the mind-body problem and mental causation*. Cambridge, MA: MIT Press.

Kim, J. (2010). *Essays in the metaphysics of mind*. New York: Oxford University Press.

Lacasse, J. R., & Leo, J. (2005). Serotonin and depression: A disconnect between the advertisements and the scientific literature. *PLoS Medicine*, *2*(12), e392.

Lebowitz, M. S., & Ahn, W. (2014). Effects of biological explanations for mental disorders on clinicians' empathy. *Proceedings of the National Academy of Sciences*, *111*(50), 17786–17790.

Lebowitz, M. S., & Ahn, W. (2015). Emphasizing malleability in the biology of depression: Durable effects on perceived agency and prognostic pessimism. *Behaviour Research and Therapy*, *71*, 125–130.

Lillard, A. S. (1996). Body or mind: Children's categorizing of pretense. *Child Development*, *67*(4), 1717–1734.

Linden, D. E. J. (2006). How psychotherapy changes the brain: The contribution of functional neuroimaging. *Molecular Psychiatry*, *11*(6), 528–538. https://doi.org/10.1038/sj.mp.4001816

Lozano, A. M. (2011). Harnessing plasticity to reset dysfunctional neurons. *New England Journal of Medicine*, *364*(14), 1367–1368. https://doi.org/10.1056/NEJMcibr1100496

Macdonald, K., Germine, L., Anderson, A., Christodoulou, J., & McGrath, L. M. (2017). Dispelling the myth: Training in education or neuroscience decreases but does not eliminate beliefs in neuromyths. *Frontiers in Psychology*, *8*, 1314. https://doi.org/10.3389/fpsyg.2017.01314

Miresco, M. J., & Kirmayer, L. J. (2006). The persistence of mind-brain dualism in psychiatric reasoning about clinical scenarios. *American Journal of Psychiatry*, *163*(5), 913–918.

Moncrieff, J. (2006). Psychiatric drug promotion and the politics of neoliberalism. *British Journal of Psychiatry*, *188*(4), 301–302.

Moncrieff, J., & Cohen, D. (2006). Do antidepressants cure or create abnormal brain states? *PLoS Medicine*, *3*(7), e240. https://doi.org/10.1371/journal.pmed.0030240

Mudrik, L., & Maoz, U. (2014). "Me & my brain": Exposing neuroscience's closet dualism. *Journal of Cognitive Neuroscience*, *27*(2), 211–221.

Munkholm, K., Paludan-Müller, A. S., & Boesen, K. (2019). Considering the methodological limitations in the evidence base of antidepressants for depression: A reanalysis of a network meta-analysis. *BMJ Open*, *9*(6), e024886.

O'Connor, C., Rees, G., & Joffe, H. (2012). Neuroscience in the public sphere. *Neuron*, *74*(2), 220–226.

Proctor, C. C. T. (2008). Clinicians' and laypeople's beliefs about the causal basis and treatment of mental disorders *[Unpublished doctoral dissertation]*. Yale University, New Haven, CT.

Robinson, H. (2017). Dualism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2017 ed.). Retrieved from https://plato.stanford.edu/archives/fall2017/entries/dualism/

Rutherford, B. R., Wager, T. D., & Roose, S. P. (2010). Expectancy and the treatment of depression: A review of experimental methodology and effects on patient outcome. *Current Psychiatry Reviews*, *6*(1), 1–10. https://doi.org/10.2174/157340010790596571

Schroder, H. S., Duda, J. M., Christensen, K., Beard, C., & Björgvinsson, T. (2020). Stressors and chemical imbalances: Beliefs about the causes of depression in an acute psychiatric treatment sample. *Journal of Affective Disorders*, *276*, 537–545. https://doi.org/10.1016/j.jad.2020.07.061

Sharfstein, S. S. (2005). Big pharma and American psychiatry: The good, the bad, and the ugly. *Psychiatric News*, *40*(16), 3–4. https://doi.org/10.1176/pn.40.16.00400003

Shtulman, A., & Lombrozo, T. (2016). Bundles of contradiction: A coexistence view of conceptual change. In D. Barner & A. S. Baron (Eds.), *Core knowledge and conceptual change* (pp. 49–67). New York: Oxford University Press.

Slingerland, E., & Chudek, M. (2011). The prevalence of mind-body dualism in early China. *Cognitive Science*, *35*, 997–1007. https://doi.org/10.1111/j.1551-6709.2011.01186.x

Smart, J. J. (1959). Sensations and brain processes. *Philosophical Review*, *68*(2), 141–156.

Sohler, N., Adams, B. G., Barnes, D. M., Cohen, G. H., Prins, S. J., & Schwartz, S. (2016). Weighing the evidence for harm from long-term treatment with antipsychotic medications: A systematic review. *American Journal of Orthopsychiatry*, *86*(5), 477–485. https://doi.org/10.1037/ort0000106

Stanovich, K. E. (1989). Implicit philosophies of mind: The dualism scale and its relation to religiosity and belief in extrasensory perception. *Journal of Psychology*, *123*(1), 5–23. https://doi.org/10.1080/00223980.1989.10542958

Stoljar, D. (2017). Physicalism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2017 ed.). Retrieved from https://plato.stanford.edu/archives/win2017/entries/physicalism/

UN Human Rights Council. (2017). Report of the Special Rapporteur on the right of everyone to the enjoyment of the highest attainable standard of physical and mental health. 28 March, A/HRC/35/21. Retrieved from https://www.refworld.org/docid/593947e14.html

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.

Weisman, K., Legare, C. H., Smith, R. E., Dzokoto, V. A., Aulino, F., Ng, E., Dulin, J. D., Ross-Zehnder, N., Brahinsky, J. D., & Luhrmann, T. M. (in press). Similarities and differences in concepts of mental life among adults and children in five cultures. *Nature Human Behaviour*.

Wellman, H. M., & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology*, *43*, 337–375. https://doi.org/10.1146/annurev.ps.43.020192.002005

Wellman, H. M., & Johnson, C. N. (2008). Developing dualism: From intuitive understanding to transcendental ideas. In A. Antonietti, A. Corradini, & J. E. Lowe (Eds.), *Psycho-physical dualism today: An interdisciplinary approach* (pp. 3–36). Lanham, MD: Lexington Books.

Westphal, J. (2016). *The mind–body problem*. Cambridge, MA: MIT Press.

Whitaker, R., & Cosgrove, L. (2015). *Psychiatry under the influence: Institutional corruption, social injury, and prescriptions for reform*. New York: Palgrave Macmillan.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supporting information