

This article was downloaded by: [Yale University Library]

On: 20 August 2013, At: 09:34

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Philosophical Psychology

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/cphp20>

### Is moral internalism supported by folk intuitions?

Caj Strandberg & Fredrik Björklund

Published online: 02 Apr 2012.

To cite this article: Caj Strandberg & Fredrik Björklund (2013) Is moral internalism supported by folk intuitions?, *Philosophical Psychology*, 26:3, 319-335, DOI: [10.1080/09515089.2012.667622](https://doi.org/10.1080/09515089.2012.667622)

To link to this article: <http://dx.doi.org/10.1080/09515089.2012.667622>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

# Is moral internalism supported by folk intuitions?

Caj Strandberg and Fredrik Björklund

*In the metaethical debate on moral internalism and externalism, appeal is constantly made to people's intuitions about the connection between moral judgments and motivation. However, internalists and externalists disagree considerably about their content. In this paper, we present an empirical study of laymen's intuitions about this connection. We found that they lend surprisingly little support to the most celebrated versions of internalism, which provides reason to be skeptical of the evidential basis for these views.*

*Keywords:* Experimental Philosophy; Externalism; Internalism; Intuition; Moral Judgment; Moral Motivation

## 1. Introduction

One of the most debated issues in contemporary metaethics concerns the connection between moral judgments and motivation. Moral internalism states that there is a necessary connection between moral judgments and motivation to act, whereas moral externalism states that no connection of this type holds.

The debate between internalists and externalists follows a certain fixed dialectic. Internalists invite us to consider thought experiments in which a person makes use of moral language without being morally motivated. They maintain that we respond to such cases by being reluctant to ascribe a moral judgment to the person unless she is motivated accordingly, which they take as evidence for their position. On the received view, the principal rejoinder available to externalists is to maintain that we can conceive of an amoralist, i.e., a person who holds a moral judgment without being morally motivated. Internalists typically reply in either of two ways. The first reply is to maintain that, properly understood, we do not find the amoralist conceivable, although it might seem so at first glance.<sup>1</sup> The second reply is to accept that the amoralist in question does constitute a counterexample to the version of

---

Caj Strandberg is a Lecturer in Practical Philosophy at the University of Gothenburg (Sweden).

Fredrik Björklund is a Professor of Psychology at Lund University (Sweden).

Correspondence to: Caj Strandberg, Department of Philosophy, Linguistics, and Theory of Science, University of Gothenburg, Box 200, SE-405 30, Gothenburg, Sweden. Email: [caj.strandberg@gu.se](mailto:caj.strandberg@gu.se)

internalism under discussion, but then propose a modified version of internalism that accommodates this type of amoralist. The latter strategy allows internalists to set up new thought experiments modeled on the amended version of internalism. Now the usual dialectic can continue with the latest version of internalism in focus.

It is generally presumed in this dialectic that a key type of evidence consists in people's pre-philosophical intuitions about the connection between moral judgments and motivation.<sup>2</sup> As we have seen, however, there is considerable controversy about what these intuitions amount to, and advocates of externalism and the various versions of internalism consequently insist that the relevant data support their particular position. Indeed, the fact that a substantial aspect of this dialectic basically seems to consist in the two camps ascribing different intuitions to people has made some commentators conclude that the internalism debate has reached a stalemate. However, despite the general appeal to intuitions, they have not been subject to systematic empirical research.<sup>3</sup> As a first attempt to rectify this fact, we carried out an empirical study in order to explore laymen's intuitive conception of the connection between moral judgments and motivation. The study suggests that the most celebrated versions of internalism find unexpectedly little support in these intuitions, which provides reason to be skeptical of the evidential basis for these views. We hope that the study will help release the internalist debate from the aforementioned stalemate by stimulating further research on people's conception of the connection between morality and motivation.

## 2. Versions of Internalism

A simple version of internalism can be characterized as follows:

*Generic Internalism:* it is necessary that, if a person judges that she is morally required to  $\phi$ , then she is, at least to some extent, motivated to  $\phi$ .

In what follows, we will understand the various versions of internalism as conceptual claims, which means that "necessary" is read as "conceptually necessary."<sup>4</sup> Thus understood, the internalist views we will consider are assumed to be true in virtue of the conventional meaning of moral terms and sentences. In metaethics, it is common to formulate conceptual claims by saying that a moral sentence expresses a certain mental state, which is referred to as "moral judgment."<sup>5</sup> On Generic Internalism, this means that a sentence of the type "I am morally required to  $\phi$ " expresses a judgment such that, if a person judges that she is morally required to  $\phi$ , then she is motivated to  $\phi$ . It follows that it would be incorrect to ascribe such a judgment to a person unless she is motivated to  $\phi$ . Corresponding considerations hold for the other versions of internalism we will consider.

It is important to observe that, in order for a version of internalism to get support from intuitions in the way described in the last section, it needs to be understood as a conceptual claim. The reason is that the kind of thought experiments internalists make use of appeal to our linguistic intuitions concerning under what circumstances it would be correct to maintain that a person holds a certain moral judgment.<sup>6</sup>

Moreover, we will understand the relevant internalist claims as stating that a person who judges that she is required to  $\phi$  only needs to be motivated to some extent to  $\phi$ , not that  $\phi$ ing is what she has strongest motivation to do. The reason is that it seems quite possible that a person genuinely judges that she is required to  $\phi$  but has stronger motivation to do something else because, say, she thinks it would further her own welfare.

Let us now turn to the amoralist objection and the manners in which internalists have responded to it by proposing amended versions of this view.<sup>7</sup> According to one version of the amoralist objection, it is conceivable that a person judges that she is required to  $\phi$  without being motivated to  $\phi$  if she suffers from a certain mental condition, such as apathy, compulsion, depression, emotional disturbance, or psychopathy (Mele, 1996; Stocker, 1979). As a consequence, most internalists opt for the following type of internalism:

*Conditional Internalism:* it is necessary that, if a person judges that she is morally required to  $\phi$ , then she is, at least to some extent, motivated to  $\phi$ , if she satisfies condition C.

There are different versions of Conditional Internalism depending on how condition C is specified. The most renowned version maintains that C consists in being *practically rational* (Korsgaard, 1996; Smith, 1994; van Roojen, 2010). According to another influential version, C consists in functioning in a *psychologically normal* manner, but there are also other alternatives (see, e.g., Blackburn, 1998, pp. 59–68; Björnsson, 2002; Dreier, 1990, pp. 12–14; Gibbard, 2003, p. 154). This type of internalism is compatible with the possibility that a person who judges that she is morally required to  $\phi$  is not motivated to  $\phi$ , namely if she fails to satisfy C. It is assumed that a person who suffers from a mental condition of the type just mentioned fails to satisfy C, which explains why she might lack motivation to  $\phi$ .

According to another version of the amoralist objection, it is conceivable that a person judges that she is required to  $\phi$  without being motivated to  $\phi$  even if she does not suffer from any mental condition of the kind referred to above (see, e.g., Nichols, 2004, chapter 3; Roskies, 2003; Svavarsdóttir, 1999). For example, it might be argued that a moral cynic may judge that she is morally required to perform a certain action without being motivated to do so and that this can be the case even if she satisfies condition C. Internalists might respond to this objection by weakening internalism even further. Recently, internalists have argued that the necessary connection between moral judgments does not hold on an individual but at a societal level (Bedke, 2009; Blackburn, 1998, pp. 59–68; Greenspan, 1998; Tresan, 2006, pp. 150–152, 2009; see also Foot, 1978, pp. 189–207):

*Communal Internalism:* it is necessary that, if a person judges that she is morally required to  $\phi$ , then she is member of a society in which people's moral judgments generally are accompanied by motivation.

There are different versions of Communal Internalism depending on how the exact connection between moral judgments and motivation in a society is specified

(Francén, 2010, pp. 119–120). This type of internalism is compatible with the possibility that a person who judges that she is required to  $\phi$  is not motivated to  $\phi$  even if she fulfils condition C. However, it is incompatible with the possibility that a person who judges that she is morally required to  $\phi$  is not part of a society in which there is a general connection between moral judgments and motivation. In support of this contention, proponents of Communal Internalism employ various thought experiments concerning the connection between moral judgments and motivation in entire societies, but they also find reasons for their view in considerations of the essentially social nature of morality (Bedke, 2009, pp. 4–7; Lenman, 1999, pp. 445–446; Tresan, 2009, pp. 185–186; see also Gert & Mele, 2005, pp. 276–278).<sup>8</sup>

### 3. An Empirical Study of Internalist Intuitions

We conducted an empirical study in order to investigate to what extent the various versions of internalism are supported by people's pre-philosophical intuitions about the connection between moral judgments and motivation.

#### 3.1. Method

**Participants.** 176 participants (58 women and 115 men, 3 failed to report sex) with a mean age of 22, 62 ( $SD = 3, 51$ ) volunteered to participate in the study. Almost all of the participants were students. 82 of them (47%) had taken about half a semester of philosophy at university, whereas the rest had not taken any philosophy. Importantly, none of the participants had gone through any courses covering internalism or externalism.

**Design, materials, and procedure.** The materials consisted of six scenarios each of which was followed by a question. Each scenario described a person, Anna, who confronts a morally relevant action but is not motivated accordingly. The six pairs of scenarios and questions were designed to test different versions of internalism.

Participants were informed that the study concerned moral motivation. They were asked to read the scenarios carefully but to respond intuitively to the succeeding questions. The scenarios were presented in three different orders, to which the participants were randomly assigned, so that we could check whether order of presentation affected the results (which it did not).

We carefully tried to formulate the scenarios and succeeding questions so as to probe the most relevant versions of internalism in ways that avoided distinct philosophical vocabulary. Four considerations are particularly worth mentioning. First, since internalist claims merely entail that a person who holds a moral judgment needs to be motivated to *some extent* to perform the action in question, we stressed, both in the scenarios and the succeeding questions, that Anna is *not motivated at all, not to any extent*, to perform the action. Moreover, in the instructions, “being motivated” to perform an action was explained as “being inclined” to perform it. In these ways, we hoped to make clear that it is not only the case that Anna does not perform the action, or that she is not most motivated to perform the action, but that

she has no inclination whatsoever to do so. Second, since internalist claims maintain a necessary connection between a person's moral judgment, in the form of a mental state as specified above, and motivation, we asked if it could be the case that Anna *thinks* that she is morally required to perform the action in question. In this way, we hoped to make clear that the question is not merely whether Anna, say, understands what it means to claim that someone is morally required to perform an action, but whether *she* actually is of the *opinion* that *she* is morally required to perform the action in question. Third, since we were interested in *conceptually necessary* versions of internalism, we asked if it *could* be the case that Anna thinks she is morally required to perform the action in question even if the relevant motivation is absent. In this way, we avoided the philosophically complicated expression "necessary" and its derivations. Fourthly, we tried to describe the relevant mental conditions (apathy, depression, and psychopathy) in ways that agree with psychological science but still are comprehensible for non-psychologists.

The most basic type of scenario was *Simple*:

*Simple*

Anna is watching a TV programme about a famine in Sudan. In the TV programme, it is shown how the starving are suffering and desperately looking for food. At the same time, Anna is not motivated at all, not to any extent, to give any money to those who are starving.

Question: could it be the case that Anna thinks she is morally required to give some of her money to the starving even if she not motivated at all to do so?

This basic type of scenario was varied so as to generate four further scenarios, which, in addition to the content in *Simple*, contained the following pieces of information:

*Normal Functioning*

Anna is mentally healthy and functions normally. For example, she is not depressed, apathetic, emotionally disturbed, psychopathic, or the like.

*Apathy*

Anna is apathetic. She is listless and generally reluctant to make any effort, even if it would serve her own interests. As a result of Anna being apathetic, she is not motivated at all, not to any extent, to give any money to those who are starving.

*Depression*

Anna is deeply depressed. Most of the time she is sad and tired. She has also difficulties concentrating and is not interested in doing the things that use to appeal to her. As a result of her depression, Anna is not motivated at all, not to any extent, to give any money to those who are starving.

*Psychopath*

Anna is a psychopath. She is incapable of empathizing with the feelings of other human beings and she often behaves in ways that other people strongly dislike. Moreover, she does not feel any remorse for her behavior towards other people. As a result of her being a psychopath, Anna is not motivated at all, not to any extent, to give any money to those who are starving.

The second type of scenario was *Community*:

*Community*

Imagine a society X that in most respects is similar to ours. The citizens of X look roughly as we do, behave roughly as we do, and like pretty much the same things as we do. Citizens of X know that there are other people who are starving and every now and then they watch on TV how these people are suffering and desperately looking for food. None of the citizens of X is ever motivated, not to any extent, to give any money to those who are starving.

Question: could it be the case that citizens of X think that they are morally required to give some of their money to the starving even if none of them is motivated to do so?

We also carried out two supplementary studies the relevant results of which are discussed in section 3.3.

3.2. Results

A statistical analysis of the extent to which the different scenarios were related to one another (reliability analysis) revealed that the participants tended to respond in the same manner across the different scenarios, or to put it more technically, that the responses to the different scenarios tended to measure the same latent construct (Cronbach alpha = 0.76). The correlation between each scenario and the average of the rest of the scenarios (item-total correlation) was the lowest for *Psychopath* ( $r = 0.30$ ), indicating that it may be a special case.

As can be seen in table 1, there were significantly more “yes” responses in all scenarios except *Psychopath*, where there were significantly more “no” responses.<sup>9</sup>

There was no effect of having studied philosophy except in *Psychopath* (Chi-2 = 4, 69,  $p = 0.03$ ). There was no difference among the participants that had studied philosophy as regards how many that responded “yes” (42) and “no” (40) to this scenario, but among the participants that had not studied philosophy, it was less common to respond “yes” (31) than “no” (58). In other words, it was the responses of those participants that had not studied philosophy that caused the reversed pattern of responses to *Psychopath*. There were no gender differences as regards any of the scenarios.

**Table 1.** Frequencies, Chi-2, and  $p$ -values for the responses to the different scenarios.

Scenario	Number of “Yes”	Number of “No”	Chi-2	$p$
Simple	133	43	46,02	0.001
Normal functioning	139	37	59,11	0.001
Apathy	105	70	7,0	0.008
Depression	139	36	60,62	0.001
Psychopath	74	101	4,17	0.041
Community	126	49	33,88	0.001



### 3.3. Discussion

The study has implications for Generic Internalism and Conditional Internalism, according to which there is a certain conceptually necessary connection between a person's moral judgments and her motivation. A majority of the participants responded in ways indicating that the relevant versions of internalism are not supported by their linguistic intuitions. Thus, these versions of internalism do not receive the support they generally are presumed to from people's intuitive conception of this connection. Furthermore, a majority of the participants responded in ways that run contrary to what they should be expected to do if these views were correct. The two points can be summarized by saying that the study provides reasons to be skeptical of the evidential basis for these versions of internalism. In the study we also found some additional reasons for this conclusion. As we shall see, it can be questioned whether this study provides any reason for being skeptical of the basis for Communal Internalism, but we conducted a supplementary study that basically confirmed the aforementioned conclusion also concerning this version of internalism.

#### 3.3.1. Generic Internalism

The study provides reasons to be skeptical of the basis for Generic Internalism. We find direct support for this view in the fact that a majority of the participants responded in the affirmative to *Simple* (76%). In other words, they seem to consider it possible that a person thinks she is morally required to perform an action even if she is not motivated to do it.

There is also indirect reason for this skepticism in the affirmative responses to *Normal Functioning* (79%), *Apathy* (60%), and *Depression* (79%). Generic Internalism entails that *any* person who thinks she is morally required to perform an action is motivated to do it. A majority seem to consider it possible for a person to think that she is morally required to perform an action even if she is not motivated to do so, and they responded in this manner regardless of whether she was reported to be functioning normally, or to be suffering from apathy or depression. Hence, this result provides further reason to question the basis of Generic Internalism.

It should be stressed that these responses do not mean that the fact that a person holds a moral judgment without being motivated accordingly is not in need of explanation. The questions that the participants were asked concern whether it is *possible* for a person to think she is morally required to perform an action even if she is not motivated to do so, not whether this would seem "peculiar" (or the like). Thus, as far as these responses go, our conception of moral judgments might be such that there is a very close connection between moral judgments and motivation, although this connection need not be conceptually necessary. Similar claims hold for the other versions of internalism we consider below. This means in turn that externalists cannot take the present study as an excuse for not trying to account for our conception of this connection. (For an externalist explanation of the responses to the



various thought experiments considered in this paper, see Strandberg, 2011, pp. 341–369.)

### 3.3.2. Conditional Internalism

The study also provides reasons to be skeptical of the basis for Conditional Internalism. First, a majority of the participants (79%) responded in the affirmative to *Normal Functioning*. They seem to consider it possible that a person thinks she is morally required to perform an action without being motivated to do it even if she functions normally. In the scenario the fact that a person “functions normally” was explained as “she is not depressed, apathetic, emotionally disturbed, psychopathic, or the like.” As we have seen, advocates of Conditional Internalism insert in the basic version of internalism a condition C, which is assumed to imply that the person in question does not suffer from any of the mentioned mental conditions. Conditional Internalism thus means that, if a person judges that she is morally required to perform an action, then she is motivated to perform it, given that she functions normally in the sense specified in this scenario. Hence, the responses give reason to doubt that laymen intuitively think that there is a conceptually necessary connection between a person’s moral judgments and her motivation even on the assumption that she fulfils C.

Second, a majority of the participants responded affirmatively to *Apathy* (60%) and *Depression* (79%). They seem to consider it possible that a person thinks she is morally required to perform an action without being motivated to perform it if she suffers from apathy or depression. The fact that a majority of the participants responded in a similar manner to, on the one hand, *Normal functioning*, and on the other hand, *Apathy* and *Depression*, provides a problem for Conditional Internalism. If this view were correct, it would be expected that most participants respond in the negative to *Normal Functioning*, but in the affirmative to *Apathy* and *Depression*. The reason is that the former scenario does not contain any explanation as to why the person in question fails to be morally motivated that could be interpreted in terms of condition C, whereas the latter two scenarios do contain such explanations. However, most participants responded affirmatively to the former scenario to at least as high an extent as to the two latter ones. Hence, they responded in a manner running contrary to what should be the case on Conditional Internalism.

Third, it is also worth noticing that the participants responded differently to the three scenarios referring to mental conditions. As regards *Depression*, a majority (79%) seem to consider it possible that a person thinks she is morally required to perform an action without being motivated to do so if she is depressed. As regards *Apathy* the majority was considerably smaller (60%), and as regards *Psychopath*, there were significantly fewer participants—indeed a minority—who thought the corresponding scenario to be possible (42%).<sup>10</sup>

The diversity in the responses to these three scenarios illustrates a particular difficulty for Conditional Internalism in specifying condition C. Advocates of this view generally agree that there are a variety of different mental conditions that can cause a person not to be motivated in accordance with her moral judgment, such as

addiction, apathy (listlessness), compulsion, depression, emotional disturbance, and psychopathy. According to Conditional Internalism, the fact that a person who judges that she is morally required to perform an action is not motivated to perform it is explained in terms of C. This means that the specification of C needs to accommodate all those, but only those, mental conditions that might cause a person not to be motivated in accordance with her moral judgment. More precisely, the specification of C needs to be such that each mental condition which can have this result instantiates a specific way of failing to satisfy C, and that there are no mental conditions which cannot have this result that instantiate a failure to satisfy C. In view of the various mental conditions that are thought to have this capacity, it can be seriously doubted whether there is a specification of C that fulfils this requirement. Moreover, the relevant version of Conditional Internalism is understood as a conceptually necessary claim, which means that the specification of C needs to comply with our conception of what can cause a person not to be motivated in accordance with her moral judgment.<sup>11</sup> Plausibly, this provides further reason to doubt that there is a way of specifying C that fulfils this requirement.

The aforementioned diversity in the responses to *Depression*, *Apathy*, and *Psychopath* provides independent reason for this doubt. The three mental conditions referred to in these scenarios constitute prototypical examples of failures of satisfying C as this condition is usually understood. For example, they constitute prototypical examples of failure in rationality and psychological normality. However, the diversity in the responses to these scenarios indicate that the participants had rather different views as to whether depression, apathy, and psychopathy, respectively, can cause a person not to be motivated in accordance with her moral judgment. As a consequence, the responses to these three scenarios illustrate the difficulty of specifying C in the required manner. That is, they illustrate the difficulty of specifying C in a way which accommodates those, but only those, mental conditions that, according to our conception, might cause a person not be motivated in accordance with her moral judgment.<sup>12</sup>

We would also like to call attention to a methodological consideration that is relevant in the present context. Taken in isolation, responses to the three scenarios referring to mental deficiencies—*Apathy*, *Depression*, or *Psychopath*—do not provide evidence either for or against the plausibility of Conditional Internalism. Assume first that a majority of the participants respond to any of these scenarios in the affirmative. That is, assume that they consider it possible that a person who suffers from the mental condition referred to in such a scenario thinks she is morally required to perform a certain action although she is not motivated to do it. This would not provide evidence against Conditional Internalism, since the reason why the participants respond in this way might be that they believe that the person fails to satisfy a certain condition C, which they consider as an explanation of why the person fails to be morally motivated. Assume next that a majority of the participants respond to any of these scenarios in the negative. That is, assume that they do not consider it possible that the person figuring in such a scenario thinks she is morally required to perform a certain action. This would not constitute any evidence for Conditional

Internalism, one reason being that the participants might respond in this way because they consider it impossible that a person with the mental condition in question thinks she is morally required to perform any action. That is, the explanation of such responses need not be that they consider it impossible that a person who fails to be motivated to perform an action thinks she is morally required to perform it. Rather, it might be that they consider it impossible that a person who suffers from the type of mental condition in question thinks she is morally required to perform any action at all. Thus, *in isolation* the responses to each of the three scenarios cannot constitute arguments that have any bearing on the plausibility of Conditional Internalism. However, as we have seen above, *comparisons* between responses to different relevant scenarios can constitute such arguments.

We would like to close the discussion of Conditional Internalism with considering two previous studies with relevance for our results as regards *Psychopath* and *Depression*. In a pioneering empirical study, Nichols (2004) examines laymen's conception of the connection between moral judgments and motivation in psychopaths. The participants were given a description of John, a psychopathic criminal who has hurt and killed people without reacting emotionally. John says he knows that it is wrong to hurt other people, but that he does not care if he does things that are wrong. The participants were then asked if John really understands that hurting others is morally wrong. A majority responded in the affirmative, which makes Nichols' question the idea that internalism is part of people's intuitive conception of the connection between moral judgments and motivation. Interestingly, our study gave a different result. Indeed, the participants' responses to *Psychopath* differ significantly from their responses to the other scenarios, this being the only scenario to which a majority did not respond in the affirmative.

Let us start by considering what conclusions can be drawn from the responses to the two scenarios. It should first be noticed that the affirmative responses to Nichols' psychopath scenario provide a reason to question the basis for Generic Internalism. (This corresponds to the indirect argument against Generic Internalism discussed above.) However, as was implied above, they can hardly constitute a reason to question the basis for Conditional Internalism. In particular, there is nothing in Nichols' psychopath scenario entailing that the person in question fulfils a relevant condition C, such as being rational. As a consequence, it might be argued that the participants responded as they did because they believed that, while John actually thinks that hurting others is morally wrong, he does not satisfy a certain condition C, which explains why he is not morally motivated. It should next be noticed that, as was also implied above, the negative responses to *Psychopath* do not provide an argument for Conditional Internalism. Together these two considerations illustrate our methodological point that it is only comparisons between the relevant scenarios that can constitute arguments that have any bearing on Conditional Internalism.

It might next be inquired why the responses to Nichols' psychopath scenario and to *Psychopath* differed. Answers to this question tend to be speculative, but it seems plausible to assume that they can be found in differences in the philosophical training of those who participated in the two studies.<sup>13</sup> There are particularly three

considerations that suggest this view. First, as regards *Apathy* and *Depression*, there was no significant difference in the responses depending on whether the participants had studied philosophy or not. A majority of our participants responded to these scenarios in the affirmative irrespective of their philosophical training. Second, as regards *Psychopath* it was mainly participants that had not studied philosophy who responded in the negative. Indeed, this was the only scenario in which this factor had any effect on the results. However, in Nichols' study, where a majority answered that the psychopath understands that hurting others is morally wrong, the participants consisted in students taking an introductory philosophy class (Nichols, personal communication, September 23, 2011). Third, it is difficult to see that the explanation of why the participants who had not studied philosophy responded to *Psychopath* in the negative would be that they thought that a person needs to be morally motivated in order to hold a moral judgment. The basic reason is that a majority of them responded to *Normal Functioning*, *Apathy*, and *Depression* in the affirmative, which suggests that they were not of this view regarding these scenarios. Moreover, there does not seem to be any relevant difference between *Psychopath* and these three scenarios, which suggests that this explanation applies as regards the first scenario but not as regards the three latter ones.

These considerations lead us to hypothesize that the participants who had not studied philosophy tended to think of psychopaths as persons who do not hold moral judgments at all. They consider it impossible that a psychopath, as described in *Psychopath*, thinks she is morally required to perform an action. However, it is not the fact that she fails to be morally motivated that makes them respond in this manner, but rather the belief that psychopaths cannot hold moral judgments at all.<sup>14</sup> It might further be queried what made the philosophically untrained participants respond in the indicated manner. Again, answers are bound to be speculative. However, one reasonable explanation seems to be that they consider it impossible for a person to hold a moral judgment if she, in the words of *Psychopath*, is "incapable of empathizing with the feelings of other human beings." This hypothesis deserves to be investigated further, since it might say something important about people's pre-philosophical conception of the capacity to form moral judgments.<sup>15</sup>

In a recent paper, Cholbi (2011) argues that empirical studies of depression suggest that this mental condition typically does not affect people's moral motivation. Especially, he cites empirical data suggesting that a person's depression does not have the capacity to eliminate her motivation to act in accordance with her moral judgments. We do not want to question this contention, but only point out that it has limited consequences for the results of the study we conducted. It should be recalled that the versions of internalism discussed in the present paper construe these views as conceptual claims. The responses to *Depression* indicate that a majority of the participants consider it as conceptually possible that a person thinks she is morally required to perform a certain action without being motivated to perform it in case she is depressed. This gives us reasons to doubt the basis for Generic Internalism as understood here. Similar claims apply to the other arguments employing *Depression* above. Moreover, these responses leave room for the view that

it is conceptually necessary that, if a person judges that she is morally required to perform an action, then she is motivated to perform it, granted that she fulfils a certain condition which excludes depression. These responses consequently give Conditional Internalists reason to insert a condition C in the basic formulation of internalism. In neither case does it matter whether it actually is the case that depression affects or eliminates motivation, since conceptual claims are under consideration. Thus, it can be admitted that there are empirical findings indicating that depression does not in fact influence moral motivation. However, this fails to show that we cannot legitimately appeal to people's intuitive conceptions of how depression might influence moral motivation in arguments for or against various conceptual versions of internalism.<sup>16</sup>

### 3.3.3. Communal Internalism

Whereas our original study provides reason to be skeptical of the basis for Generic and Conditional Internalism, it might be argued that it does not have similar implications for Communal Internalism. A majority of the participants in this study responded in the affirmative to *Community* (72%). Thus, they seem to consider it possible that there is an entire society in which people think that they are required to perform a certain type of action, such as giving money to the starving, but where none of them is motivated to actually perform such actions.

A reviewer of an earlier version of this paper pointed out that it can be doubted whether this result has any implications for Communal Internalism. We described Communal Internalism as the view that it is necessary that, if a person judges that she is required to perform an action, then she is member of a society in which people's moral judgments generally are accompanied by motivation. According to a reasonable interpretation of this view, the connection between people's moral judgments and motivation in such a community need not concern their moral judgments about each particular type of action.<sup>17</sup> Thus, it leaves open the possibility that people in this community think that they are required to perform a certain type of action, such as giving money to starving people, but none of them is motivated to perform that particular type of action. What this view does not allow is that this holds for all types of actions they think they are morally required to perform. As a result, it might plausibly be questioned whether the responses to *Community* have any bearing on this version of internalism.

In order to test Communal Internalism in a way that makes justice to this view, we carried out a supplementary study focusing exclusively on this version of internalism. The materials consisted of an amended version of *Community*, which was followed by a question similar to the one in the original study.<sup>18</sup> The new scenario described a society in which the citizens categorize certain actions as being morally required and other actions as not being morally required. It was reported that none of them is ever motivated, not to any extent, to perform any of the actions that, according to their classification, are morally required. The participants were then asked whether it could be the case that people in this society actually think that they are morally required to perform these actions even if they are not motivated at all to do so. Importantly, this

scenario did not include any essential reference to a particular type of action, such as giving money to starving people. As a consequence, this scenario is not open to the same objection as the original version of *Community* and it seems plausible to assume that it tests the relevant version of internalism.

We received basically the same result as regards the original version of *Community*. A significant majority of the participants (81%) responded in the affirmative. Thus, they seem to regard it as possible that there is an entire society in which people think that there are certain actions that they are morally required to perform, but where none of them is motivated at all to perform any of these actions. This can be taken to suggest that the participants did not regard it conceptually necessary that, in order for a person to think that she is morally required to perform an action, she needs to be part of a society in which people generally are motivated to perform the actions they think they are morally required to perform. As a consequence, the supplementary study provides reason to be skeptical of the basis for Communal Internalism.

In closing, we readily admit that all the empirical studies we carried out have certain methodological limitations. Let us highlight two of them. First, it is difficult to devise scenarios and questions in such ways that they accurately capture the relevant philosophical concepts and at the same time are fully understandable for laymen. Although we think we have succeeded in avoiding the most apparent difficulties of this kind in the manners explained in section 3.1, it might still be maintained that certain of the crucial concepts were not made clear enough to some of the participants, which might have influenced their responses. For example, it might be argued that the term ‘could’, as used in our questions to the participants, might be interpreted in other ways than we intended. Second, our studies do not probe all versions of internalism, although they capture those that have been most discussed in the debate. However, we consider the present studies as a starting point of a series of investigations projected to systematically examine people’s conception of the connection between morality and various action-guiding attitudes. We consequently look forward to further investigations that eliminate potential shortcomings of the present studies by, for example, explaining the relevant concepts in more detail, employing more elaborated and varied scenarios, asking a variety of different questions, and using control questions to secure that the participants completely have understood the relevant concepts.

#### 4. Conclusion

The study we conducted suggests that a conceptual form of Generic Internalism is not generally supported by laymen intuitions, since a majority of the participants seem to regard it possible that a person thinks she is morally required to perform an action without being motivated accordingly. Likewise, the study suggests that the conceptual forms of Conditional Internalism proposed in the literature are not generally supported by laymen intuitions either, since a majority of the participants responded in a similar manner as regards these claims. Moreover, a majority of the



participants responded in ways that run contrary to what they should be expected to do on the assumption that any of these internalist claims were correct. In a subsequent study, we received similar results concerning Communal Internalism. Taken together, the results of the study provide reasons to doubt the evidential basis for claiming that there is a non-trivial conceptually necessary connection between moral judgments and motivation. We want again to stress, however, that more research is needed to arrive at an empirically informed account of people's conception of the connection between moral judgments and motivation.

## Acknowledgments

We are indebted to Jon Tresan and three anonymous referees for *Philosophical Psychology* for detailed and tremendously helpful comments on an earlier version of this paper. We are also very grateful to Gunnar Björnsson, Linus Broström, John Eriksson, Claudia Jansen, Erik Janson, Mats Johansson, Petter Johansson, Shaun Nichols, Ragnar Francén Olinder, Michael Ridge, and members of the research seminar in practical philosophy at the University of Gothenburg for valuable discussions on how to investigate internalism and externalism empirically. This research was supported by the Swedish Research Council (grant number 2009–1517).

## Notes

- [1] This response is often combined with an explanation of why amoralists might seem conceivable. For two different suggestions, see Hare (1952, pp. 124–126 & 163–165) and Smith (1994, pp. 68–71).
- [2] Some authors, however, have proposed arguments that are not based on intuitions; see, e.g., Smith (1994, pp. 71–77), Svavarsdóttir (1999), and Zangwill (2008).
- [3] In a pioneering study, Nichols (2004, chapter 4) discusses people's intuitions of the connection between moral judgments and motivation in psychopaths. We discuss these results in section 3.3.2 (for a related study, see Prinz, 2007, p. 99.) Moreover, there are studies of the connection between moral judgments and motivation that draw on empirical research concerning psychopaths and patients with brain damages; see Nichols (2004, chapter 1), Cholbi (2006), Kennett and Fine (2008a), and Roskies (2003, 2006, 2008). For discussions about the possibility of testing internalism empirically, see Kauppinen (2008), Kennett and Fine (2008b), and Smith (2008).
- [4] For an overview of different forms of internalism, see Strandberg (2011, pp. 342–347), and Björklund, Björnsson, Eriksson, Francén Olinder, and Strandberg (2012). Non-conceptual versions of internalism are considered by Bedke (2009), Björnsson (2002), Mele (1996), and Nichols (2004, chapter 5).
- [5] The conceptual claim that a moral sentence expresses a certain moral judgment, in the form of a mental state, can in turn be interpreted in various ways (for one understanding of this conception, see Ridge, 2003). Moreover, the main point below can be translated into a vocabulary according to which the meaning of sentences is understood in some other way than in terms of what mental states they express. Thus, the exact understanding of this issue is presumably not crucial to our arguments. We take “moral judgment” to be neutral as regards cognitivism and non-cognitivism. We occasionally use the phrase “hold a moral judgment” (in analogy with “hold a moral belief”) to refer to a person being in the



aforementioned mental state. We prefer this phrase to “make a moral judgment,” since the latter might be thought to refer to a speech act rather than a mental state.

- [6] It is a matter of controversy to what extent it is possible to study linguistic intuitions empirically; see, e.g., Kauppinen (2007), Nadelhoffer and Nahmias (2007), and Rey (1983).
- [7] As already noted, some internalists respond to this kind of objection by denying that amorlists are conceivable. In what follows, we will consider attempts to avoid amoralist objections by defending weaker versions of internalism.
- [8] Unfortunately, our study does not do justice to all the ingenious thought experiments presented by advocates of this view.
- [9] We found the same pattern of results in a supplementary study where participants ( $N = 79$ ) rated the same scenarios on a scale from 1 (*completely disagree*) to 9 (*completely agree*). The responses to all scenarios except *Psychopath* were significantly higher than the neutral midpoint of the scale, and accordingly closer to “agree,” whereas the responses to this scenario were significantly lower than the midpoint of the scale (as tested with one-sample  $t$ -tests). All nine scenarios were administered to all participants (within-subjects design) but in four different orders. There was no indication that order affected the responses. All the scenarios and subsequent questions were presented in Swedish and are available by request from the first author.
- [10] McNemar tests reveal that significantly more participants considered it possible in *Apathy* than *Psychopath*, and in *Depression* than *Apathy* ( $ps < 0.001$ ).
- [11] More precisely, on this view a sentence of the type “I am morally required to  $\phi$ ” expresses a judgment such that, if a person judges that she is morally required to  $\phi$ , then she is motivated to  $\phi$  given that she fulfils  $C$ . Accordingly, it is part of our knowledge of the meaning of such a sentence that  $C$  accommodates those, but only those, mental conditions that might cause a person not to be motivated in accordance with her moral judgment.
- [12] It might be objected that condition  $C$  can be understood as a negation of the disjunction of all types of mental conditions that can cause a person not to be motivated in accordance with her moral judgment. However, this ad hoc solution would hardly constitute a genuine version of internalism. It might further be objected that advocates of Conditional Internalism can simply refuse to take a stance as to whether the mental conditions referred to in the argument above constitute examples of failures to satisfy  $C$ . For example, they might stay neutral as to whether these states of mind make up examples of practical irrationality. In that case, advocates of Conditional Internalism would not need to specify  $C$  in a way that accommodates these mental conditions. However, these conditions are generally assumed to constitute prototypical examples of factors that might disrupt the connection between moral judgments and motivation. Indeed, if it is not clear that these mental conditions can have this function, it can legitimately be asked which can. As a result, this move risks robbing Conditional Internalism of much of its explanatory value.
- [13] There are significant differences between Nichols’ study and ours. For instance, whereas Nichols asked whether John (the psychopath) really *understands* that hurting others is wrong, we asked whether it could be the case that Anna (in case she is psychopath) *thinks* she is required to perform a certain action. The former formulation can be interpreted to ask whether John is aware that it is generally considered to be wrong to hurt others, whereas the latter formulation does not lend itself to this interpretation to the same extent. As a result, participants might be more willing to answer the first question in the affirmative than to respond to the second question in this manner. However, this does not explain why the responses to the two psychopath scenarios differ in the particular manner described in the main text.
- [14] A person might fail to be motivated in accordance with her moral judgment more or less generally. For example, she might entirely lack the capacity for moral motivation, or she might lack moral motivation for a certain while. It might be argued that the difference in responses to, on the one hand, *Psychopath*, and, on the other hand, *Apathy* and *Depression*,

depend on the fact that people who suffer from psychopathy are considered to belong to the former category, whereas people who suffer from apathy or depression are considered to belong to the latter. This consideration deserves further discussion. However, it fails to explain the difference between the responses to *Psychopath* and the responses to Nichols' psychopath scenario.

- [15] If this hypothesis turns out to be viable, it suggests that, on people's pre-philosophical conception, the capacity to empathize with other human beings is fundamental for being able to form moral judgments. For philosophically relevant discussions on the relation between moral judgments and motivation in psychopaths, see e.g., Cholbi (2006), Kennett (2002), Kennett and Fine (2008a), Nichols (2004, chapters 1, 3, & 5), and Roskies (2003, 2006, 2008).
- [16] In Cholbi's (2011) view, internalists and externalists should stop appealing to their modal intuitions, since it can be suspected that these are shaped by antecedent theoretical commitments. Unfortunately, we cannot discuss this general methodological issue in the present paper.
- [17] We are grateful to an anonymous reviewer and to Jon Tresan for stressing this point.
- [18] We employed the following modified version of *Community*: Imagine a society X that in most respect is similar to ours: the citizens of X look roughly as we do, behave roughly as we do, and like pretty much the same things as we do. Like us, the citizens of X group actions into different moral categories. They put the label "morally right" on certain actions and the label "morally wrong" on certain other actions. In a similar manner, they classify some actions as being morally required to perform and other actions as not being morally required to perform. Their classification of various actions into moral categories is very similar to ours. For example, the action of giving money to starving people is classified as an action they are morally required to perform. However, none of the citizens of X is ever motivated, not to any extent, to perform any of the actions that, according to their classification, are morally required. *Question*: could it be the case that the citizens of X actually think that they are morally required to perform these actions, even if they are not motivated at all to do so? Participants were 101 students (42 males, 59 females, mean age 22, 46,  $SD = 2$ , 67), 5 of which had taken at least one philosophy class. The distribution of responses was very similar to that of the original community scenario; 82 responded "yes" and 19 "no" ( $\chi^2 = 39$ , 3,  $p < 0.001$ ).

## References

- Bedke, M.S. (2009). Moral judgment purposivism: Saving internalism from amorality. *Philosophical Studies*, 144, 189–209.
- Björklund, F., Björnsson, G., Eriksson, J., Francén Olinder, R., & Strandberg, C. (2012). Recent Work on motivational internalism. *Analysis*, 72, 124–137.
- Björnsson, G. (2002). How emotivism survives immoralists, irrationality, and depression. *Southern Journal of Philosophy*, 40, 327–344.
- Blackburn, S. (1998). *Ruling passions*. Oxford: Clarendon Press.
- Cholbi, M. (2006). Belief attribution and the falsification of motive internalism. *Philosophical Psychology*, 19, 607–616.
- Cholbi, M. (2011). Depression, listlessness, and moral motivation. *Ratio*, 24, 28–45.
- Dreier, J. (1990). Internalism and speaker relativism. *Ethics*, 101, 6–26.
- Foot, P. (1978). Approval and disapproval. In her *Virtues and vices* (pp. 189–207). Berkeley, CA: University of California Press.
- Francén, R. (2010). Moral motivation pluralism. *Journal of Ethics*, 14, 117–148.
- Gert, J., & Mele, A. (2005). Lenman on externalism and amorality: An interplanetary exploration. *Philosophia*, 32, 275–283.

- Gibbard, A. (2003). *Thinking how to live*. Cambridge, MA: Harvard University Press.
- Greenspan, P. (1998). Moral responses and moral theory: Socially-based externalist ethics. *Journal of Ethics*, 2, 103–122.
- Hare, R.M. (1952). *The language of morals*. Oxford: Clarendon Press.
- Kauppinen, A. (2007). The rise and fall of experimental philosophy. *Philosophical Explorations*, 10, 95–118.
- Kauppinen, A. (2008). Moral internalism and the brain. *Social Theory and Practice*, 34, 1–24.
- Kennett, J. (2002). Autism, empathy and moral agency. *The Philosophical Quarterly*, 52, 340–357.
- Kennett, J., & Fine, C. (2008a). Internalism and the evidence from psychopaths and “acquired sociopaths.” In W. Sinnott-Armstrong (Ed.), *Moral psychology: Vol. 3* (pp. 173–190). Cambridge, MA: MIT Press.
- Kennett, J., & Fine, C. (2008b). Could there be an empirical test for internalism? In W. Sinnott-Armstrong (Ed.), *Moral psychology: Vol. 3* (pp. 217–226). Cambridge, MA: MIT Press.
- Korsgaard, C. (1996). Skepticism about practical reason. In her *Creating the kingdom of ends* (pp. 311–334). Cambridge: Cambridge University Press.
- Lenman, J. (1999). The externalist and the amoralist. *Philosophia*, 27, 441–457.
- Mele, A. (1996). Internalist moral cognitivism and listlessness. *Ethics*, 106, 727–753.
- Nadelhoffer, T., & Nahmias, E. (2007). The past and future of experimental philosophy. *Philosophical Explorations*, 10, 123–149.
- Nichols, S. (2004). *Sentimental rules: On the natural foundations of moral judgment*. Oxford: Oxford University Press.
- Prinz, J. (2007). *The emotional construction of morals*. Oxford: Oxford University Press.
- Rey, G. (1983). Concepts and stereotypes. *Cognition*, 15, 237–262.
- Ridge, M. (2003). Non-Cognitivist pragmatics and Stevenson’s “Do so as well!”. *Canadian Journal of Philosophy*, 4, 563–574.
- Roskies, A. (2003). Are ethical judgments intrinsically motivational? Lessons from “acquired sociopathy.” *Philosophical Psychology*, 16, 51–66.
- Roskies, A. (2006). Patients with ventromedial frontal damage have moral beliefs. *Philosophical Psychology*, 19, 617–627.
- Roskies, A. (2008). Internalism and the evidence from pathology. In W. Sinnott-Armstrong (Ed.), *Moral psychology: Vol. 3* (pp. 191–206). Cambridge, MA: MIT Press.
- Smith, M. (1994). *The moral problem*. Oxford: Blackwell.
- Smith, M. (2008). The truth about internalism. In W. Sinnott-Armstrong (Ed.), *Moral psychology: Vol. 3* (pp. 207–216). Cambridge, MA: MIT Press.
- Stocker, M. (1979). Desiring the bad: An essay in moral psychology. *The Journal of Philosophy*, 76, 738–753.
- Strandberg, C. (2011). The pragmatics of moral motivation. *Journal of Ethics*, 15, 341–369.
- Svavarsdóttir, S. (1999). Moral cognitivism and motivation. *The Philosophical Review*, 108, 161–219.
- Tresan, J. (2006). De dicto internalist cognitivism. *Noûs*, 40, 143–165.
- Tresan, J. (2009). The challenge of communal internalism. *The Journal of Value Inquiry*, 43, 179–199.
- van Roojen, M. (2010). Moral rationalism and rationalist amoralism. *Ethics*, 120, 495–525.
- Zangwill, N. (2008). The indifference argument. *Philosophical Studies*, 138, 91–124.