Synthese (2014) 191:2445–2470
DOI 10.1007/s11229-014-0434-1

On the rationality of pluralistic ignorance

Jens Christian Bjerring · Jens Ulrik Hansen · Nikolaj Jang Lee Linding Pedersen

Received: 15 January 2014 / Accepted: 15 January 2014 / Published online: 25 March 2014

Abstract Pluralistic ignorance is a socio-psychological phenomenon that involves a systematic discrepancy between people's private beliefs and public behavior in certain social contexts. Recently, pluralistic ignorance has gained increased attention in formal and social epistemology. But to get clear on what precisely a formal and social epistemological account of pluralistic ignorance should look like, we need answers to at least the following two questions: What exactly is the phenomenon of pluralistic ignorance? And can the phenomenon arise among perfectly rational agents? In this paper, we propose answers to both these questions. First, we characterize different versions of pluralistic ignorance and define the version that we claim most adequately captures the examples cited as paradigmatic cases of pluralistic ignorance in the literature. In doing so, we will stress certain key epistemic and social interactive aspects of the phenomenon. Second, given our characterization of pluralistic ignorance, we argue that the phenomenon can indeed arise in groups of perfectly rational agents. This, in turn, ensures that the tools of formal epistemology can be fully utilized to reason about pluralistic ignorance.

J. C. Bjerring
Department of Culture and Society, Research Program for Philosophy, Aarhus University, Aarhus, Denmark
e-mail: jcbjerring@gmail.com

J. U. Hansen (✉)
Department of Philosophy, Lund University, Lund, Sweden
e-mail: jens_ulrik.hansen@fil.lu.se

N. J. L. L. Pedersen
Underwood International College, Yonsei University, Seoul, South Korea
e-mail: nikolaj@yonsei.ac.kr

Keywords Pluralistic ignorance · Epistemic rationality · Social behavior · Rational interaction · Private beliefs · Public beliefs

1 Introduction

In everyday life, we often find ourselves pondering whether to go along with the opinion of a group to which we belong, or whether to come forth with our own personal opinions. We might wonder whether to continue cheering with the group when Denmark scores against Italy, or whether to go public with our actual support for the Italians. Discrepancies between one's private beliefs and one's public behavior are well-documented in the social-psychological literature, and they play a central role in the phenomenon of *pluralistic ignorance*. Roughly put, a social situation is a situation of pluralistic ignorance when a group of individuals all have the same attitude towards some proposition or norm, all act contrary to this attitude, and all wrongly believe that everyone else in the group has a certain conflicting attitude to the proposition or norm.

The study of pluralistic ignorance emerged in social psychology, but in recent years the phenomenon has also attracted attention from formal and social epistemologists.[1] To develop a precise formal and social epistemological model of pluralistic ignorance, however, there are at least two crucial questions that need to be addressed:

1. What exactly is the phenomenon of pluralistic ignorance?
2. Can the phenomenon arise among perfectly rational agents?

Answering the first question is obviously a prerequisite for any systematic study—be it formal or conceptual—of pluralistic ignorance. While there has been considerable disagreement in the literature about how exactly to characterize pluralistic ignorance, our preferred characterization will focus on its epistemic and social interactive nature. This characterization, in turn, will bring pluralistic ignorance within the scope of formal and social epistemological investigations.

In order to fully utilize the tools of formal epistemology—such as logic and game theory but also Bayesian approaches and theories of belief revision—we need to ensure that pluralistic ignorance is a rational phenomenon. In the socio-psychological literature, people tend to think that situations of pluralistic ignorance arise because of some kind of *error*. But if the error is one of rationality—that is, if agents in situations of pluralistic ignorance fail to be rational—this would seem to constitute a serious obstacle for using logic and game theory to model pluralistic ignorance. For agents in logic and game theory are standardly taken to be subject to substantial rationality constraints, and if agents in situations of pluralistic ignorance do not meet these constraints, the worry is that the tools of logic and game theory cannot be used to reason about the phenomenon. So the second question is both interesting and important. Eventually, we argue that the question should be answered in the affirmative: it *is* possible for pluralistic ignorance to arise among perfectly rational agents. The error involved in

[1] The Copenhagen Lund workshop series on social epistemology, for instance, has taken pluralistic ignorance as one of its four focus cases (Hendricks et al. 2012).

situations of pluralistic ignorance, if there is one, is not an error of rationality. In turn, our arguments will pave the way for a proper formal treatment of pluralistic ignorance.

In Sect. 2, we shall see that pluralistic ignorance is a widespread phenomenon. It occurs in many social contexts ranging from college drinking to business organization. There is little doubt in social psychology about the potentially harmful side effects of pluralistic ignorance. Yet, there is a surprising lack of consensus about how exactly to characterize the phenomenon.

In Sect. 3, we discuss and compare various proposed characterizations of pluralistic ignorance and argue that it matters which characterization we settle for. In particular, we shall argue that certain characterizations of pluralistic ignorance are faulty in the sense that they describe some social situations as situations of pluralistic ignorance when they plausibly are not. These considerations, in turn, will motivate a normative claim about what kind of characterization should occupy center stage in studies of pluralistic ignorance. We argue that a proper characterization of pluralistic ignorance should make reference to certain distinctive *social interactive* and *epistemic* features of the social situation. More specifically, we argue that pluralistic ignorance concerning a proposition P can occur in situations with the social interactive feature that all involved individuals act contrary to their private beliefs about P and with the epistemic feature that everyone believes that the actions of others truly reflect what they believe about P.

Given our characterization of pluralistic ignorance, we discuss in Sect. 4 whether the phenomenon can arise and persist in groups of perfectly rational agents. When the notion of rationality is properly understood, we argue, there is no principled reason to doubt the rationality of pluralistic ignorance.

Finally, in Sect. 5, we conclude and discuss some of the bearings that our characterization of pluralistic ignorance has on formal studies of the phenomenon.

2 The phenomenon of pluralistic ignorance

The term "pluralistic ignorance" was coined in 1931 by the social psychologist Floyd H. Allport and his student Daniel Katz (see Katz and Allport 1931).[2] Allport's interest in pluralistic ignorance grew out of his work on "the illusion of universality of opinions"—a phenomenon that describes people's tendency to wrongly believe that their opinions are universally shared by members of a social group (O'Gorman 1986; Halbesleben and Buckley 2004). Katz and Allport (1931) reported on findings from an extensive study of students, which showed that while each individual student did not have any objection to minorities being admitted to fraternities and dormitories, each student also seemingly believed that other students might object to such admissions (O'Gorman 1986; Halbesleben and Buckley 2004). At the same time, Richard L. Schanck did an extensive study on the attitudes of residents in a small isolated village with a strong Methodist presence. On some issues, he found that there was

[2] Although there are examples of phenomena that resemble pluralistic ignorance in the literature before the work of Allport and Katz, they are not referred to by the term "pluralistic ignorance" (O'Gorman 1986). For more on the genealogy of the term, see O'Gorman (1986) and Halbesleben and Buckley (2004).

a strong discrepancy between the residents' private and public attitudes (O'Gorman 1986; Halbesleben and Buckley 2004).

The social-psychological literature provides many examples of pluralistic ignorance. We will mention three key examples to give the reader a good sense of what our target phenomenon is:

Classroom case
A teacher has just finished presenting some difficult material in class and asks the students whether they have any questions. Although each student does not fully understand the material, no one asks a question. Based on the observation that no student in the class asks a question, each student believes that everyone but him believes that the material was not difficult. To avoid being publicly displayed as the only one who did not understand the material, no student dares to ask a question.[3]

College drinking case
A group of freshmen students have just arrived at their new dorm. At the inauguration party, each student drinks excessively, although each student in fact believes that drinking is not enjoyable. Upon observing the excessive drinking of others, however, each student forms the belief that everyone but him believes that drinking is enjoyable. To avoid being publicly displayed as the boring one, every student continues to drink excessively at the party.[4]

Emperor's case
In Hans Christian Andersen's fable, "The Emperor's New Clothes" (1837), we meet two impostors who sell imaginary clothes to an emperor. They claim that those who cannot see the clothes are either not fit for their office or just truly stupid. Fearing that others will consider him unfit for his office and truly stupid, the emperor—as well as anyone else—pretends to be able to see the garment. Yet, everyone believes that the emperor is in fact naked. Based on the observation that everyone acts as if the emperor is dressed, however, each person forms the belief that everyone but him believes that the emperor is dressed. To avoid being publicly labelled as someone who is unfit for his office or truly stupid, everyone pretends that the emperor is dressed—except for the little boy who after a while cries out: "but the emperor has nothing on at all!"[5]

The above cases serve to highlight one of the main characteristics of pluralistic ignorance: the discrepancy between the private beliefs of agents in a social group and

[3] Miller and McFarland (1987) have studied and documented this type of pluralistic ignorance in classrooms. Miller and McFarland (1991) claims that "[…] pluralistic ignorance occurs in this situation because students believe that fear of embarrassment is influencing their behavior more than the behavior of the other students" (Miller and McFarland 1991, p. 298).

[4] This case is inspired by the extensive studies in social psychology of the drinking habits among college students and the influence that peers have on these habits (Borsari and Carey 2001). Prentice and Miller (1993) report on four studies of alcohol consumptions among students at Princeton, each of which documented the presence of the following phenomenon on campus: the students believed that the average student was much more comfortable with alcohol than they were themselves. In this sense, it was documented that the alcohol norm on campus differed considerably from the students' private attitudes about alcohol consumption.

[5] As we shall see in due course, this fictional case is useful for making certain conceptual points about the nature of pluralistic ignorance.

the public belief seemingly supported by the group's behavior.[6] In the next section, we will make the characterization of pluralistic ignorance more precise.

There are many well-documented reasons in the socio-psychological literature for why pluralistic ignorance arises and persists in social groups. Prentice and Miller (1993) mention two. First, they mention the "differential interpretation hypothesis". In cases such as the College Drinking Case, this hypothesis suggests that students misleadingly act as if they are more comfortable with drinking than they actually are, and that they fail to interpret the (drinking) behavior of others as misleading. Second, they advance the "differential encoding hypothesis". In cases such as the College Drinking Case, this hypothesis suggests that students fail to realize that their own actions appear much more in favor of drinking that their private beliefs are. In Halbesleben et al. (2007), two additional causes of pluralistic ignorance are discussed, namely "minority influence" and "desire to maintain social identity". Minority influence appears when a majority of people in a group takes the view of a minority to be representative of the whole group.[7] A desire to maintain social identity may lead to pluralistic ignorance because people in a group subscribe publicly to a view that they do not privately endorse in order to maintain social affiliation with the group. In the language of game theory, people may choose to subscribe publicly to a view that they do not privately endorse in order to coordinate with the social group. And they may do so because they can be characterized as playing a coordination game, in which the payoffs associated with coordinating with the group are higher than those associated with acting on their private beliefs.

These examples of what may cause pluralistic ignorance to arise point to an important correlation between the behavior of the group and the way this behavior is interpreted by an individual in the group. We will return to this correlation in more detail in Sect. 3.3.

[6] For the purposes of this paper, we characterize situations of pluralistic ignorance in terms of a discrepancy between private and public *beliefs*. But note that other attitudes could in principle be used to characterize pluralistic ignorance. For instance, several characterizations of pluralistic ignorance involve attitudes towards *norms* rather than beliefs about propositions. According to Centola et al. (2005), "[p]luralistic ignorance describes a situation where a majority of group members privately reject a norm, but assume (incorrectly) that most others accept it" (Centola et al. 2005, p. 1010). As suggested by footnote 4, the College Drinking Case can naturally be reformulated as involving norms rather than beliefs. However, since nothing substantial in our discussion hangs on the difference between "having a belief about a proposition" and "having an attitude towards a norm", we will work with the first locution in what follows.

[7] According to Centola et al. (2005), a minority can also cause pluralistic ignorance in a social network with the right structure through *informational cascading* effects—a concept first introduced by Bikhchandani et al. (1992). In Centola et al. (2005), pluralistic ignorance is modeled using agent-based models—more specifically, using two-dimensional cellular automata representing the social network. They make the implicit assumption that agents form beliefs about other agents' private attitudes towards a given norm by observing whether the neighboring agents comply with the norm and whether they enforce it. Based on such observations, the agents themselves decide whether to comply with the norm and enforce it. Insofar as a minority of agents strongly support a perhaps highly unpopular norm, their model shows how such a norm can spread throughout the network in cascading effects—much like chain reactions—and result in convergence on the unpopular norm. Note that Miller and McFarland (1991) denies that cases of minority influence lead to cases of pluralistic ignorance in general: "Pluralistic ignorance is a state of uniqueness that arises when individuals misinterpret the similar behavior of similar others, not when they generalize inappropriately from the dissimilar behavior of dissimilar (unrepresentative) others" (Miller and McFarland 1991, p. 296).

It is also well-documented in the socio-psychological literature that pluralistic ignorance may often have bad consequences. Generally, which consequences may result from situations of pluralistic ignorance seem to depend on how agents choose to act in such situations. Prentice and Miller (1993) mention three strategies that agents can adopt in cases of pluralistic ignorance to reduce the discrepancy between private and public beliefs. First, agents can bring their private beliefs closer to the belief that is seemingly supported by the group's behavior. Second, agents can try to bring the group's belief closer to their own private beliefs. Third, agents can reject or alienate themselves from the social group.[8]

Prentice and Miller classify the first option as the most simple way to dissolve the conflict between private and public beliefs:

> Given that the last two of these options will often appear too costly or too difficult to effect, at least in the short run, the simplest way for individuals to eliminate the discrepancy is to change their private attitudes. (Prentice and Miller 1993, p. 244)

In singling out this option as the most simple one, Prentice and Miller seem to factor in two sorts of considerations: cost and ease of implementation. Cost immediately springs to mind when considering the possibility of adopting the third strategy: if a person alienates or isolates herself from a given social group, this may sever potentially valuable ties at the social level and cause a feeling of discomfort at the psychological level. Difficulty of implementation springs to mind in connection with the second strategy: it may simply be too difficult for a person to change public opinion so it conforms with her own private opinion. In addition, it may be too costly for her to do so. For in making explicit the divergence between her private opinion and the public opinion, she may be rejected from the social group and face the public embarrassment potentially associated with holding a minority view.

But there are also reasons to believe that the first strategy is associated with bad effects or costs. First, if too many agents in a social group adopt the first strategy, the unsupported behavior of the group may well perpetuate. For example, if each student in the College Drinking Case aligns her behavior with that of the group, students are likely to continue drinking more than they feel comfortable doing. Given that this kind of drinking behavior may have several bad consequences, adopting the first strategy may have too. Second, if too many agents in a social group adopt the first strategy, the risk of creating a *bystander effect* increases (cf. Prentice and Miller 1993, p. 254). Traditional examples of bystander effects involve cases where individuals in a group do not offer help in a situation where help is in fact needed. To illustrate, consider a group of agents who witness a situation of emergency but without acting. Each agent observes the behavior of the other agents in the group to determine whether intervention is necessary. Because every agent takes the lack of intervention from others as a sign that no help is needed—or that the situation is not an emergency situation—the end result is that no one intervenes and offers their assistance. Since the first strategy above

[8] In their study, Prentice and Miller (1993) found evidence that male students adopted the first strategy and internalized the drinking norm seemingly supported by the group, whereas female students adopted the third strategy and alienated themselves from the group.

effectively recommends aligning one's private beliefs and behavior with the behavior of the group, it is clear how this strategy can increase the risk of bystander effects—and hence how this strategy can have potentially harmful side effects.

So irrespective of how agents react in situations of pluralistic ignorance, the phenomenon is associated with potentially harmful side effects or costs. Interestingly, these potentially bad consequences also crop up in the context of organizations.[9] Halbesleben et al. (2007), for instance, argue that pluralistic ignorance may increase employees' negative feelings that they are different from others in the organization, and that it may eschew employees from sharing their actual opinions on a topic with colleagues in order to maintain group identity. In turn, these consequences may cause higher stress and lower degree of commitment among employees. For the organization as a whole, pluralistic ignorance may then cause a weak organizational culture that is not actually supported by its members, and it may result in poor decision-making because the employees do not express their private beliefs about the optimal procedures.

So the examples and potentially bad side effects of pluralistic ignorance are numerous. However, as we shall see in the next section, numerous are also the different characterizations of pluralistic ignorance that people have proposed.

3 Characterizing pluralistic ignorance

There seems to be no consensus in the literature about how to define "pluralistic ignorance". What is more, to the knowledge of the authors, there is currently no analysis of the different definitions and their mutual relations. In this section, we will present a number of different characterizations suggested by the literature and compare them. In light of this comparison, we will then present what we take to be the most adequate definition of pluralistic ignorance.

3.1 Different definitions of pluralistic ignorance

The first central question is whether to regard pluralistic ignorance as a phenomenon tied entirely to the psychology of an individual. On one approach, pluralistic ignorance involves a cognitive or psychological error made by an individual in a social situation—for instance, an error of social comparison. In line with this view, we find Halbesleben and Buckley (2004) who state that "[p]luralistic ignorance is a social comparison error where an individual holds an opinion, but mistakenly believes that others hold the opposite opinion." This characterization makes pluralistic ignorance come out as a phenomenon of an individualistic nature because it is tied to the psychological state of a single agent.

[9] For an overview of the ramifications of pluralistic ignorance in management and organizations, see Halbesleben et al. (2007) and Halbesleben and Buckley (2004). An example of pluralistic ignorance in organization theory is found in Westphal and Bednar (2005) where, "[…] under conditions of low performance, there may be a systematic tendency for outside directors to underestimate the extent to which fellow directors share their concerns about the viability of the firm's corporate strategy" (Westphal and Bednar 2005, p. 262).

Prentice and Miller (1993) define pluralistic ignorance as "[…] a psychological state characterized by the belief that one's private attitudes and judgments are different from those of others, even though one's public behavior is identical" (Prentice and Miller 1993, p. 244).[10] As such, Prentice and Miller also take pluralistic ignorance to be a psycho-individualistic phenomenon, but not purely so, insofar as they also put weight on the role of behavior. O'Gorman (1986) claims that Allport also viewed pluralistic ignorance as a psychological phenomenon:

> To begin with, Allport never clearly developed the critical distinction between the psychological distortions of individual perception and the mistaken ideas shared by people. As a rule, he considered shared error as a special case of individual distortions. (O'Gorman 1986, p. 338).

In particular, O'Gorman argues that Allport believed that only individuals are real, and that he never seriously considered the possibility that social factors are characteristic of a phenomenon such as pluralistic ignorance.

Let us turn this psycho-individualistic characterization of pluralistic ignorance into a definition:

> (PI_1) *"Pluralistic ignorance" refers to a situation where an individual member of a group believes some proposition P and mistakenly believes that the other members of the group believe $\neg P$.*

Reflection on the literature, however, lends support to the thought that pluralistic ignorance is not adequately characterized as a phenomenon grounded entirely in the psychological state of an *individual* agent. Although Halbesleben and Buckley (2004) and Prentice and Miller (1993) explicitly define pluralistic ignorance in terms of an individual agent's cognitive error, they all mainly discuss situations where *all* individuals in a group make the same type of cognitive error. So on a charitable reading we might interpret them as claiming that paradigmatic situations of pluralistic ignorance concern situations in which *all* members of the group make a cognitive error. This kind of idea is made explicit by Krech and Crutchfield who define pluralistic ignorance as a situation where "[…] *no one believes, but everyone believes that everyone else believes*" (Krech and Crutchfield 1948, pp. 388–389).

We can capture the idea that pluralistic ignorance is a situation where all agents in a social group make the same cognitive error as follows:

> (PI_2) *"Pluralistic ignorance" refers to a situation where a group of individuals all believe some proposition P and they all mistakenly believe that the other members of the group believe $\neg P$.*

Although (PI_2) highlights the social aspect of pluralist ignorance, it still leaves the phenomenon reducible to purely psychological states of individuals.

Several authors, however, claim that pluralistic ignorance goes beyond mere psychology and involves irreducible social components. O'Gorman (1986) is one of the strongest advocates of this view. He says:

[10] Prentice and Miller have adopted this definition from Miller and McFarland (1991) who state a very similar definition.

> This is not to suggest that individual cognitive or perceptual distortion is irrelevant to understanding collective misconceptions. But it does serve to illustrate that a theory of pluralistic ignorance must consider the conditions that lead individuals, on the one hand, to impute to others ideas and actions different from their own and, on the other hand, to act contrary to their own beliefs and values (O'Gorman 1986, pp. 334–335).

(PI_1) and (PI_2) both emphasize the discrepancy between private and public attitudes in situations of pluralistic ignorance. O'Gorman's point in the above passage is that this is not enough: to adequately characterize pluralistic ignorance we must also try to explain why the discrepancy is there.

O'Gorman goes on to say that the core question in the study of pluralistic ignorance is this: "to what extent is a social environment, primarily its culture and social structure, so arranged as to mislead its inhabitants about its major characteristics?" (O'Gorman 1986, p. 335). This question presupposes that the social context plays a crucial role in explaining the discrepancy between the private and public spheres. We agree. Inspired by O'Gorman's comments, let us give a third definition of pluralistic ignorance that captures this social aspect:

> (PI_3) *"Pluralistic ignorance" refers to a situation where a group of individuals (i) all believe some proposition P, (ii) all mistakenly believe that the other members of the group believe $\neg P$, and (iii) the situation described by (i) and (ii) is brought about by the group being in a special kind of social context.*

Although we took the characterization of pluralistic ignorance in Prentice and Miller (1993) to be a version of either (PI_1) or (PI_2), there is evidence that they are aware of the importance of the social context. For instance, they claim that pluralistic ignorance "develops most commonly under circumstances in which there is a widespread misrepresentation of private views. In these cases, people's tendency to rely on the public behavior of others to identify the norm leads them astray" (Prentice and Miller 1993, p. 225). In this sense, it is also possible to view their characterization of pluralistic ignorance as lying somewhere between (PI_1), (PI_2), and (PI_3).

As stated, note that (PI_3) is merely schematic. Different versions will result from different specifications of what constitutes the "special kind of social context". There are two lines we can take on this matter. On the one hand, we may hold there is an exact definition of what counts as the relevant kind of social context, and, as such, that (PI_3) will eventually be replaced by a definite characterization. On the other hand, we may hold that there is a range of social contexts that are equally legitimate or relevant when it comes to the characterization of pluralistic ignorance. We take the first of these two lines and suggest, in Sect. 3.3, that the relevant kind of social context is one in which agents base their individual beliefs about other agents' beliefs on observation of their behavior.

3.2 Distinguishing pluralistic ignorance

While it is clear that a type (PI_3) situation implies a type (PI_2) situation, and that a type (PI_2) situation implies a type (PI_1) situation of pluralistic ignorance, we will argue in

the next two sections that these implications do not run in the other direction. We focus first on (PI_1) and (PI_2) and argue that they fail to adequately characterize pluralistic ignorance because they *overgenerate* cases of pluralistic ignorance. That is, (PI_1) and (PI_2) classify certain social situations as situations of pluralistic ignorance, although they plausibly are not. In light of this criticism, we suggest in Sect. 3.3 that a specific version of (PI_3) should be the focal point in investigations of pluralistic ignorance.

We first give two examples of social situations in which (PI_1) is satisfied but (PI_2) and (PI_3) are not. In turn, we use these examples to argue that (PI_1) overgenerates cases of pluralistic ignorance.

Example 1 You are a scholar of mathematical logic and have just arrived as a guest researcher at a university abroad. You have seen an announcement of an interesting math seminar and decide to attend it. Being new to place, you arrive well in time for the seminar and decide to sit in the back of the seminar room. As a mathematical logician, you believe that philosophy is interesting, but you are also well aware that mathematicians often do not share this belief. Since you are attending a math seminar, you hence form the belief that other people in the seminar room believe that philosophy is not interesting. Unbeknownst to you, however, it turns out that you by accident have joined the weekly philosophy seminar, where all the other participants are well-acquainted and know of each other that they all like philosophy. So you believe that philosophy is interesting but also erroneously that everyone else believes that philosophy is not interesting.

Example 2 You are teaching a course on social psychology and today's lecture is on racial segregation. Personally, you believe that racial segregation is wrong, but during your lecture you are shocked to find that most of your students seem to believe that racial segregation is not wrong. Unbeknownst to you, however, the students deliberated before class and decided to play a trick on you—all of them actually believe that racial segregation is wrong.

For (PI_1) to be satisfied, it is only required that a single individual believes P and erroneously believes that everyone else believes $\neg P$. These requirements are met in Examples 1 and 2, and so, by the lights of (PI_1), these examples qualify as cases of pluralistic ignorance. In contrast, for (PI_2) and (PI_3) to be satisfied, it is required that all individuals mistakenly believe that the other individuals believe $\neg P$. This requirement is not met in Example 1 and 2, and so, by the lights of (PI_2) and (PI_3), these examples do not qualify as cases of pluralistic ignorance. It follows that (PI_1) is less restrictive than both (PI_2) and (PI_3).

Now let us turn to the issue of adequacy. If (PI_1) adequately characterized pluralistic ignorance, Examples 1 and 2 should count as situations of pluralistic ignorance. But we think that they should not:

In Example 1, you are mistaken about the identity of the social group, and based on this mistake, you form a false belief about their beliefs about philosophy. Yet, your mistake in Example 1 is minor and easily explainable in simple epistemological terms, and in contrast to the paradigmatic cases of pluralistic ignorance in the literature, your mistake is not shared by other people in the social group. Nor does Example 1 seem to involve any of the special psychological phenomena or cognitive errors that we

mentioned in Sect. 2. In fact, it is hard to see that Example 1 should be associated with any of the potentially harmful side effects that are normally associated with situations of pluralistic ignorance. If this is true, there are good reasons to hold that Example 1 does not describe a situation of pluralistic ignorance, and hence that (PI_1) does not adequately characterize the phenomenon.

Similar observations apply to Example 2. In Example 2, you are simply the victim of a practical joke that tricks you into forming a false belief about the students' beliefs about racial segregation. Your mistake in Example 2 can be explained using familiar concepts from the epistemological literature on testimony. Also, none of the psychological phenomena, cognitive errors, or potentially harmful side effects relevant to pluralistic ignorance seem to be involved in Example 2. So again it seems as if (PI_1) does not adequately characterize pluralistic ignorance. Rather, (PI_1) overgenerates situations of pluralistic ignorance.

On the face of it, (PI_2) and (PI_3) seem to fare better than (PI_1). For while Examples 1 and 2 qualify as situations of pluralistic ignorance according to (PI_1), they do not according to (PI_2) and (PI_3). But, as we shall argue now, (PI_2) is still not adequate. To that end, we give two examples of social situations in which (PI_2)—and therefore (PI_1)—is satisfied but (PI_3) is not. In turn, we use these examples to argue that (PI_2) overgenerates cases of pluralistic ignorance.

Example 3 50 freshmen—none of whom know each other—arrive at their new campus dorm. Each freshman finds in his room an introduction package that includes several invitations to dorm parties that promise large supplies of free alcohol. While each freshman believes that drinking is not enjoyable, he forms—based on the information in the introduction package—the belief that the other freshmen believe that drinking is enjoyable.

Example 4 Consider a group of 50 young Danes in the age 16–19 who are randomly picked and do not know each other. Assume further that each person in the group does not have a Facebook profile and believes that Facebook is a waste of time. If each person in this group, however, is asked whether they believe that the other people in group have a Facebook profile, they will most likely answer "yes". For they are well aware of the widespread use of Facebook among other young people. In turn, they will most likely also believe that the other people in the group believe that Facebook is not a waste of time.[11]

For (PI_2) to be satisfied, it is required that each individual in the social group believes P and erroneously believes that everyone else in the group believes $\neg P$. Since these requirements are met in Examples 3 and 4, they describe situations in which (PI_2) is satisfied. On the other hand, (PI_3) is not satisfied, as we shall show in Sect. 3.3. Given this, we can hold that (PI_2) is less restrictive than (PI_3).

Now let us turn to the issue of adequacy. If (PI_2) provided an adequate characterization of pluralistic ignorance, Examples 3 and 4 should count as situations of pluralistic ignorance. But we think that they should not:

[11] According to Statistics Denmark, 89% of Danes in the age 16–19 had a Facebook profile in 2010 (Danmarks Statistik 2011, p. 26, Tabel 7).

First, as above, there are relatively straightforward epistemological explanations of why agents believe what they do in the two examples. In Example 3, they invoke relevant background information about the stereotypical picture of dorm parties in addition to the basic information from the introductory package. In Example 4, they invoke relevant background knowledge about the high popularity of Facebook in the young Danish population. Second, none of the psychological phenomena and cognitive errors normally associated with situations of pluralistic ignorance seem operative in the two examples. In Example 3, the individual agents do not act as if they are more comfortable with drinking than they actually are, they do not misinterpret the behavior of others, and they do not obviously face any social identity problems pertaining to group membership and the like—of course, all these things might change *when* the actual dorm parties take place. Similarly, in Example 4, the individual agents do not act as if they are Facebook fans, they do not misinterpret the behavior of others in the group, and they do not face any social pressure or identity problems pertaining to group membership. Finally, it is not obvious that Examples 3 and 4 have any of the potentially harmful side effects that are normally associated with situations of pluralistic ignorance. There is, for instance, no need to expect that Examples 3 and 4 should lead to bystander effects. For a bystander effect to occur, direct observational interaction among members in the social group is needed. But in Examples 3 and 4, there is no such interaction. If this is true, there are good reasons to hold that Example 3 and 4 do not describe situations of pluralistic ignorance, and hence that (PI_2) overgenerates cases of pluralistic ignorance.

In contrast to the examples above, paradigmatic cases of pluralistic ignorance involve some explanation of how people in the relevant social context *interact*. In particular, observations of other people's behavior play a crucial role. In the Classroom Case, agents observe that the other agents in the group do not ask any questions about the class material. Based on this observation—or based on this interaction with the social group—they form the false belief that everyone else in the group believes that the material is easy. In the College Drinking Case, agents observe that the other agents in the group drink excessively at the dorm party. Based on this observation, they form the false belief that everyone else in the group believes that drinking is enjoyable. In the Emperor's Case, agents observe that the other agents in the group act as if the emperor is dressed. Based on this observation, they form the false belief that everyone else in the group believes that the emperor is dressed. Similarly, in arguing why Examples 3 and 4 do not plausibly count as situations of pluralistic ignorance, the lack of observational interaction among agents in the relevant social group plays a crucial role. In the next section, we propose a definition of pluralistic ignorance that is sensitive to this kind of interaction among agents.

3.3 The epistemic and social interactive characterization of pluralistic ignorance

Examples 1 to 4 show that (PI_1) and (PI_2) overgenerate situations of pluralistic ignorance. To overcome these shortcomings, we must enrich the definition of pluralistic ignorance to disqualify Examples 1 to 4 from counting as such. To do so, we will complete the schematic definition (PI_3).

In order to complete (PI_3), we need to specify what counts as a "special kind of social context". Our considerations from above point towards two features that are distinctive of situations of pluralistic ignorance. First, agents in such situations act contrary to their private beliefs. Second, they form their false beliefs about other people's beliefs on the basis of observations of their actions. O'Gorman (1986) and Miller and McFarland (1991) are in broad agreement. O'Gorman, as we have already seen, says that an account of pluralistic ignorance must take into account

> "[…] the conditions that lead individuals, on the one hand, to impute to others ideas and actions different from their own and, on the other hand, to act contrary to their own beliefs and values" (O'Gorman 1986, pp. 334–335).

In a similar vain, Miller and McFarland describe pluralistic ignorance as a situation

> "[…] in which people take a public position on a social issue that misrepresents their private position. Pluralistic ignorance arises here because individuals erroneously perceive the public expressions of others to represent faithfully their private positions" (Miller and McFarland 1991, p. 292).

Let us first take a closer look at the second distinction feature of pluralistic ignorance. Bicchieri (2006) makes the point that "[…] common to all contexts in which pluralistic ignorance occurs is the lack of transparent communication (or of any communication at all) among individuals" (Bicchieri 2006, p. 187). But notice that this lack of (transparent) communication does not exclude the flow of information in situations of pluralistic ignorance. On the contrary, people retrieve extensive amount of information from *observing* other people's behavior in situations of pluralistic ignorance. Miller and McFarland (1991) argue that mutual observability often is predominant among agents in such situations. In fact, they argue that pluralistic ignorance often arises in social groups precisely because of this mutual observability (Miller and McFarland 1991, pp. 290–291). Likewise, in the paradigmatic cases of pluralistic ignorance, people form their (false) beliefs about other people's beliefs about P based on *observations* of their behavior in the group—as opposed to through direct or explicit communication. In the College Drinking Case, each student comes to believe erroneously that every other student in the social group believes that drinking is enjoyable by observing the excessive drinking they all engage in. Something similar happens in the Classroom Case and the Emperor's Case: agents form their beliefs about the beliefs of other agents in the group by observing how they behave.

Based on these reflections on our key examples, we can isolate a distinctively *epistemic* component of pluralistic ignorance: people use behavior as guide to what other people believe. That is, they take the actions of others as strong evidence for their private beliefs. There is nothing epistemically peculiar about this belief-forming method; we use it all the time. Behavior is typically a good, yet fallible guide to figuring out what people believe. And paradigmatic cases of pluralistic ignorance make salient this fallibility because they are characterized precisely by being situations in which behavior is *not* a good guide to belief. What explains this? The short answer is "social interactions within the group". But let us provide a more elaborate answer—one that will make it clear that pluralistic ignorance has a genuinely social component.

Forming beliefs about other people's beliefs by observing their behavior is a form of social interaction.[12] On its own, however, this kind of interaction does not lead to pluralistic ignorance. All people in the social group must also act contrary to their private beliefs. In this sense, the "special kind of social context" that occurs in situations of pluralistic ignorance is one in which people, on the one hand, form their beliefs about others based on observations of their behavior, and, on the other hand, act contrary to their private beliefs. In line with O'Gorman's view, then, pluralistic ignorance is a genuine social phenomenon. (PI_3) is superior to the other definitions of the phenomenon exactly because it mentions this social aspect.

Given that pluralistic ignorance is a genuine social-epistemological phenomenon, and bearing in mind that (PI_3) is only schematic, we propose the following version of (PI_3) as an adequate definition of pluralistic ignorance:

(PI_4) *"Pluralistic ignorance" refers to a situation where the individual members of a group*

(i) *all privately believe some proposition* P;
(ii) *all believe that everyone else believes* $\neg P$;
(iii) *all act contrary to their private belief that* P *(i.e. act as if they believe* $\neg P$*)*; *and where*
(iv) *all take the actions of the others as strong evidence for their private beliefs about* P.[13]

Whereas conditions (i) and (ii) in (PI_4) are familiar from (PI_2) and (PI_3), conditions (iii) and (iv) capture the two distinctive features of pluralistic ignorance that we elucidated above.

In contrast to (PI_1) and (PI_2), (PI_4) does not characterize the situations described in Examples 1 to 4 as situations of pluralistic ignorance. Examples 1 and 2 are not cases of pluralistic ignorance according to (PI_4) for the same reasons that they are not cases of pluralistic ignorance according to (PI_2). In Example 3, the freshmen do not interact with each other and they do not form their erroneous higher-order beliefs by observing how the other students act. Similarly, in Example 4, the young Danes do not interact with each other and they do not observe each other's behavior either. Hence neither Example 3 nor Example 4 are situations of pluralistic ignorance according to (PI_4).[14]

[12] As Miller and McFarland (1991) say, "the victims of pluralistic ignorance are not observers, they are participants in the group dynamic who know that their own behavior belies their internal state and cannot be taken at face value" (Miller and McFarland 1991, p. 297).

[13] We assume that the meaning of the phrases "acting as if P" and "acting contrary to one's belief about P" are sufficiently clear. We are aware, however, that on some occasions it can be non-trivial to determine whether an action is for or against a given proposition. An agent may, for instance, find it difficult to decode by observation which action an agent is seemingly performing, or he may misinterpret the other agent's action. Nevertheless, we will bracket such issues since our main points do not rely on these finer details. Like us, note that (Bicchieri, 2006, ch. 5) also stresses the importance of conditions similar to (iii) and (iv) in discussing the emergence of pluralistic ignorance. Unlike us, however, Bicchieri does not engage in a detailed discussion of the proper definition of pluralistic ignorance, and eventually she adopts a definition similar to the one in Prentice and Miller (1993).

[14] Of course, we may go on to extend the stories in Example 3 and 4 in various ways to include interaction and observations among the freshmen and the young Danes. And such extensions could lead to cases of

Before moving on, it is useful to look a little closer at the relationship between the four conditions in (PI_4). As shown by Examples 3 and 4, condition (iii) does not follow from (i) and (ii). Adding (iv) to (i) and (ii) does not result in (iii) being satisfied either, as the individuals might act in ways that do not give any indications as to whether they believe P or not. Condition (iv) does not follow from the other conditions either. For individuals can believe P and form the relevant beliefs in (ii) without basing these beliefs on observations of the group's behavior—they might form them by guessing, say. One might think that (ii) is superfluous by thinking that it follows from (iii) and (iv). But this is not the case. For although people might take the actions of others as strong evidence for their private beliefs about P, they might nevertheless take seriously the possibility that others act insincerely. As such, they might refrain from relying on observations of other people's behavior to form their beliefs about others.

Let's take stock. In this section, we have provided reasons for thinking that (PI_4) does a good job of capturing pluralistic ignorance, and that it nicely captures the idea that pluralistic ignorance is a genuine social-epistemological phenomenon. While it is obvious that (i) and (ii) are epistemic components of pluralistic ignorance, our discussion shows that (iv) is another key epistemic component. Conditions (iii) and (iv) in (PI_4) also highlight the importance of the social interaction among agents in situations of pluralistic ignorance, and jointly, they make it clear that pluralistic ignorance cannot be reduced to mere psychological states of the individuals in the relevant social group.

Given our characterization of pluralistic ignorance, we will in the next section argue that the phenomenon can arise in groups of rational agents—henceforth, if nothing else is mentioned, "pluralistic ignorance" refers to type (PI_4) situations of pluralistic ignorance. As mentioned in the introduction, considerations on rationality are important for understanding the nature of pluralistic ignorance and for understanding in what ways it might or might not prove possible to model the phenomenon using well-known formal frameworks such as logic and game theory. In particular, if we can show that no error of rationality need be involved in situations of pluralistic ignorance, we will have paved the way for formal investigations of the phenomenon.

4 Is pluralistic ignorance a rational phenomenon?

In arguing that pluralistic ignorance can arise in groups of rational agents, the task is to show that it can be rational for agents to conform to conditions (iii) and (iv) in (PI_4). No such task, it seems, is involved with respect to (i) and (ii). On the one hand, it is obvious that it can be rational for agents to conform to (i) by having, say, excellent reasons for their private belief in P. On the other hand, whether or not it is rational for agents to conform to (ii) seems derivative on whether or not it is rational for them to conform to (iii) and (iv). If it is rational for an agent to rely on this observational evidence that everyone in his group acts as if $\neg P$—cf. (iii)—and to take the actions

Footnote 14 continued
pluralistic ignorance—Example 3 could turn into the College Drinking Case, for instance—but they need not.

of others as strong evidence for their private beliefs about P—cf. (iv)—then it is also rational for the agent to form the belief in (ii) that everyone but him believes $\neg P$. Given that (i) is unproblematic, we hence only need to show that it can be rational for agents to conform to (iii) and (iv) in order to show that pluralistic ignorance can arise among rational agents.

To do so, we proceed as follows: in Sect. 4.1, we clarify the relevant notion of rationality. In Sect. 4.2, we discuss some considerations that support the thought that some errors of rationality are involved in cases of pluralistic ignorance. In Sect. 4.3, we then go on to argue that these considerations in fact do not support this thought, and that pluralistic ignorance can arise in groups of rational agents.

4.1 Rationality

The notion of rationality on which we rely is consequentialist in nature. The standing of actions and beliefs vis-à-vis rationality is determined by the goodness of their consequences. In turn, the degree to which the consequences of a given action or belief are good is determined by the degree to which they promote things of value.

It is standard to distinguish between two kinds of rationality: *pragmatic rationality* and *epistemic rationality*.[15] Both kinds of rationality can be understood along consequentialist lines. Pragmatically rational actions and beliefs are ones that promote things of pragmatic value (or have good pragmatic consequences), and epistemically rational actions and beliefs are those that promote things of epistemic value (or have good epistemic consequences). This, of course, is merely schematic until it is specified what things are of pragmatic and epistemic value respectively.

The most widely held view on epistemic value is a certain form of monism—what might aptly be labeled 'veritic monism'. According to this view, truth is the only non-derivative epistemic good. Any other epistemic good is only derivatively valuable—that is, valuable in virtue of bearing some significant relationship to truth. For instance, one might say that justification is only derivatively valuable because what makes justification epistemically valuable is the fact that justified belief is more likely to be *true*. Here we stay with the philosophical mainstream and adopt veritic monism.

But what about pragmatic value? An action or belief is pragmatically valuable to the extent that it promotes practical interests or goals. Sometimes 'pragmatic value' is also used as a catch-all phrase to denote any kind of value that is not epistemic. Against the background of veritic monism, this way of understanding pragmatic value would have any value other than truth and its derivatives come out as pragmatic. It does not matter for present purposes which of these two characterizations we adopt: the things we assume are pragmatically valuable below count as such according to both characterizations.

In light of the distinction between pragmatic and epistemic value and rationality, a number of interrelated points or issues rise to the surface. First, beliefs and actions can possess two different types of value, and each type marks a dimension of normative evaluation. That is, beliefs and actions can qualify as more or less rational—along the

[15] For more on the distinction between pragmatic and epistemic rationality, see, e.g., Christensen (2004).

pragmatic as well as the epistemic dimension—depending on the extent to which they promote the relevant kind of value. This point immediately raises a second issue: is it possible to compare the two kinds of value and rationality, or weigh them against each another? This would seem necessary in order to speak generally of the *overall* value or rationality of a given belief or action—or of its value or rationality, *all things considered.*

The kinds of cases that call for a weighing or comparison of values are the ones where the two types of value come into conflict—where the target belief or action does poorly with respect to one, but well with respect to the other. Suppose, for example, that Bill Gates told you that he will give you $10,000,000 if you believe that $2 + 2 = 5$. In that case believing that $2 + 2 = 5$ would—if possible—be of great pragmatic value to you. Just think of all the wonderful things that you could do with that kind of money. On the other hand, holding the belief would be of no epistemic value to you. Indeed, if anything, having the belief would lead to negative value because it is false, and you know that this is so. In cases like this it is not clear what to say about value and rationality all things considered, at least not without specifying a measure for weighing and comparing the pragmatic and epistemic dimensions of evaluation.

It is notoriously difficult to come up with a principled measure for weighing and comparing different types of goods or values. Indeed, some might even think that this challenge cannot be met: there are some beliefs and actions for which no verdict can be reached about whether they are valuable or rational all things considered. Thankfully, we do not have to take a stance on this difficult and controversial matter here. In arguing that pluralistic ignorance *can* arise and persist in groups of perfectly rational agents, we do have in mind rationality all things considered—and henceforth, if nothing else is indicated, we mean all-things-considered rationality by "rationality". However, since we are advancing a *possibility* claim, we are not required to show that an all-things-considered verdict can be reached in all cases. Strictly speaking, one will suffice.

4.2 The irrationality of pluralistic ignorance

Given a characterization of rationality, we consider now the question of whether pluralistic ignorance is a rational phenomenon. We consider in this section first arguments purporting to show that it is not, and then in the next section arguments purporting to show that it is. Ultimately, we will be arguing that situations of pluralistic ignorance *can* indeed arise among rational agents.

To make a case for the irrationality of pluralistic ignorance, we may attempt to argue that the phenomenon cannot arise in groups of rational agents. Consider an agent who, unbeknownst to himself, is in a situation of pluralistic ignorance as characterized by (PI_4):

(**PI-Rat**)
(a) I believe P.
(b) I believe that everyone in my context—including me—acts as if $\neg P$.
(c) Because everyone—including me—acts as if $\neg P$, I believe that everyone—except me—believes that $\neg P$.

There are two issues that seemingly cause trouble for the thought that pluralistic ignorance can persist in this situation. First, if the agent in the situation described by (PI-Rat) indeed believes P, it would seem *epistemically* rational for him to act on this belief and reflect it in behavior. After all, behaving in ways that directly contradict what one believes is not conducive for promoting truth: either the belief will fail to represent what is the case, or one's actions will not align with what is the case. If so, the agent ought to dissolve the situation of pluralistic ignorance by ceasing to act as if $\neg P$ when he believes P. Second, if the agent in (PI-Rat) indeed believes (a) to (c), it would seem *epistemically* rational for him to use this evidence to dissolve the situation of pluralistic ignorance. For given that he has already entertained (a) to (c), he is a short step from entertaining the following:

(**PI-Rat**) (cont.)

(d) But if I act as if $\neg P$, even though I believe P, why could other people in my context not experience the same discrepancy between their private beliefs and their public behavior?

(d_1) But if (d), I cannot rule out the possibility that there are other people in my context who actually believe P, although they act as if $\neg P$. If I cannot rule out this possibility, however, I am not justified in maintaining my higher-order belief in (c). So I should not continue believing that everyone else—except me—believes that $\neg P$.

If the reasoning in (d) and (d_1) in fact dissolves the situation of pluralistic ignorance, and if the error-possibility in (d_1) is taken seriously by the agent, he ought epistemically dissolve such situations from mere reflection on his available evidence.

Of course, we might appeal to various *pragmatic* values in order to explain why agents continue to act as if $\neg P$, although they in fact believe P. Presumably, the pragmatic values associated with various psychological and sociological factors such as fear of losing face and of losing social identity can help explain the relevant differences between an agent's private beliefs and public behavior. Similarly, there are presumably various pragmatic values that can help explain why agents ignore—or at least assign a very low credence to—the possibility mentioned in (d), i.e. that other agents in the group act as if $\neg P$, although they in fact believe P. Prentice and Miller, for instance, say that agents in situations of pluralistic ignorance

> […] typically make the mistake of assuming that even though others are acting similarly, they are feeling differently. Their own behavior may be driven by social pressure, but they assume that other people's identical behavior is an accurate reflection of their true feelings (Prentice and Miller 1993, p. 244).

We expect that well-known empirical phenomena from social psychology such as attributions errors and various cognitive biases can help explain why agents in situations of pluralistic ignorance "feel" that they are different from other agents, though they act similarly and in fact share the same beliefs.

Various psychological and sociological factors may go a long way in grounding the pragmatic values that help explain why agents in fact do not dissolve situations of pluralistic ignorance. Yet, given the considerations above, such explanations are

compatible with the idea that (all-things-considered) *rational* agents in fact ought dissolve such situations by mere reflection. For when we specify the content of P and describe in detail the relevant situation of pluralistic ignorance, we might be in a position to argue that the epistemic values overall outweigh the pragmatic ones, and hence that it would be rational, all things considered, for the agent to dissolve the situation of pluralistic ignorance. Examples are controversial, but perhaps an argument could be made that the epistemic values of going public with one's private belief in the Classroom Case—when appropriately specified, as we shall see later—can outweigh the pragmatic values of staying quiet. If so, it is compatible with various psychological and sociological explanations of pluralistic ignorance that only agents that fail to be rational, all things considered, can end up in situations of pluralistic ignorance.

4.3 The rationality of pluralistic ignorance

Despite the reasons for believing in the irrationality of pluralistic ignorance, we believe that there are also strong reasons to believe in its rationality. To argue for this, we now present a three-part argument. The first part makes a case for the claim that it may be pragmatically rational for agents to continue to conform to condition (iii) of (PI_4) in situations of pluralistic ignorance. The second part supports the claim that it may be epistemically rational for agents to continue to conform to condition (iv) of (PI_4). The third part extends these conclusions from respectively pragmatic and epistemic rationality to rationality all things considered.

First part. For the first part of the argument that concerns the pragmitic rationality of conforming to (iii), the central question is this: how can agents have reasons to act as if $\neg P$, although they in fact believe P? They can, it seems, whenever the pragmatic reasons for coordinating their behavior with the group outweigh those of coordinating their behavior with their private beliefs. While such situations are familiar from game theory where games such as stag hunting show that the benefits of coordinating one's behavior with the group outweigh those of not doing so, they also seem present in most paradigmatic cases of pluralistic ignorance.

For notice that pragmatically bad repercussions result from diverging from the group's behavior in the standard situations of pluralistic ignorance. In the College Drinking Case, for instance, the cost of drinking excessively may be substantially smaller than the potential cost of getting isolated from one's peers. And similarly, praising the Emperor's (imaginary) clothes seems considerably less costly than running the risk of being stamped as truly stupid or unfit for one's office. Since matters such as social exclusion and social identity can impact people's lives in significant ways, it is thus not hard to see that individual agents may often be pragmatically rational in coordinating their behavior with the group's behavior. In turn, we have an argument showing that it *may* be *pragmatically* rational for agents to conform to (iii) in (PI_4).

Second part. For the second part of the argument that concerns the epistemic rationality of conforming to (iv), the question is this: which epistemic factors can help explain why an agent ignores—or at least assigns a very low credence to—the possibility in (d) that other agents in the social group, like him, do not reflect their private beliefs in their public behavior?

As argued above, the agent's main reason for maintaining the relevant higher-order belief in (c) stems from observational evidence: he observes that other agents behave as if $\neg P$. Simultaneously, however, he realizes upon introspective reflection—as described in (d) and (d_1)—that other agents, just like him, may act as if $\neg P$ although they in fact believe P. So the central question is whether the agent's observational evidence outweighs his introspective evidence. If it does, we have an explanation of why the agent in (PI-Rat) has epistemic reasons to ignore the possibility that the actions of others do not truly reflect what they believe.

In most cases of pluralistic ignorance, it seems, an agent has indeed good epistemic reasons to give more weight to his observational evidence than to his introspective evidence. In standard cases such as the College Drinking Case, agents lack any observational evidence for seriously doubting that the group's behavior does not reflect what each member in the group in fact believes. In the College Drinking Case, there is no striking conflict between the agent's belief that drinking is not enjoyable, his observations of the group's behavior, and his higher-order belief that everyone but him finds drinking enjoyable. Rather, the agent's observations of the excessive drinking in the group makes it epistemically rational for him to maintain the higher-order belief in question. If so, it is sensible to hold that the agent's observational evidence outweighs his introspective evidence in these sorts of cases, and hence that it remains epistemically rational for him to ignore the possibility that the actions of others do not reflect their private beliefs. In turn, we have an argument showing that it *may* be *epistemically* rational for agents to conform to (iv) in (PI_4).

Notice, however, that there might be certain specific situations of pluralistic ignorance in which it is not obviously epistemically rational for an agent to take the actions of others as strong evidence for their private beliefs. If an agent finds himself in a case similar to the one described in the Emperor's Case, for instance, there is an intuition that he should pay less attention to the observational evidence of the group's behavior than in cases such as the College Drinking Case—bracketing, for the purpose of making a conceptual point, that the former is only a fictional tale. Intuitively, agents in the Emperor's Case have less epistemic reasons to maintain the relevant higher-order belief in (c) than they have in the College Drinking Case. In the former case, there is perfect perceptual evidence for holding that the emperor is naked, and given that perceptual evidence is a highly reliable belief-forming mechanism, there also seems to be very good epistemic reasons for believing that the emperor is indeed naked. Upon observing that others in the group seemingly do not believe that the emperor is naked, an enlightened agent may reason as follows:

> Unless I am radically deceived about the nature of the external world, there is something fishy going on in the current situation. For normally, my fellow people do trust their senses and make qualified predictions about the world. So unless they have suddenly started to believe that our current situation is one in which appearances systematically fail to reflect reality, the fact that they act as if the emperor is not naked cannot truly reflect what they actually believe.

Insofar as the agent has good epistemic reasons for believing that the emperor is naked but only few, if any, epistemic reasons for believing that he is in an evil-demon

world—and that other agents believe that they are in such a world—it seems that he has only few epistemic reasons for maintaining the higher-order belief that everyone but him believes that the emperor is naked. If this is true, the agent's introspective evidence may outweigh his observational evidence in certain situations. And if it does, it is not epistemically rational for him to ignore—or to assign a low credence to—the possibility that the actions of others do not reflect what they in fact believe. Derivatively, there are reasons to hold that pluralistic ignorance *can* be dissolved purely by rational reflection in certain specific situations.

In situations of pluralistic ignorance, we have so far argued that it *may* be pragmatically rational for agents to coordinate their behavior with the social group, and that it *may* be epistemically rational for agents to believe that the observed behavior of the social group reflects what each member in fact believes. So while various psychological and sociological errors might be involved in all situations of pluralistic ignorance, no error of pragmatic rationality need be involved with respect to (iii) and no error of epistemic rationality need be involved with respect to (iv).

Third part. For the third part of the argument, the central question is whether the considerations above put us in a position to conclude that situations of pluralistic ignorance need not involve any errors of rationality *all things considered*. While we have shown that it might be *pragmatically* and *epistemically* rational to conform to (iii) and (iv) respectively, we still have not shown that it is rational all things considered to do so.

Given the comments in Sect. 4.1, however, we know which conditions need be met for it to be rational all things considered to conform to (iii) and (iv). For it is rational all things considered to conform to (iii) when pragmatic considerations in favor of conforming to (iii) trump or outweigh epistemic considerations in favor of not doing so, and rational all things considered to conform to (iv) when epistemic considerations in favor of conforming to (iv) trump or outweigh pragmatic considerations in favor of not doing so. Since we have already argued that these conditions can be met in standard cases of pluralistic ignorance, we thus have an argument showing that it *may* be *rational all things considered* for agents to conform to both (iii) and (iv) in (PI_4).

Accordingly, we can say that pluralistic ignorance may arise in groups of rational (all things considered) agents when the following two conditions—in conjunction with the other conditions in (PI_4)—are met:

> (**Con 1**) It is pragmatically rational for agents to coordinate their behavior with the social group and the pragmatic advantages of doing so outweigh the epistemic disadvantages of doing so; and
> (**Con 2**) it is epistemically rational for agents to believe that the observed group behavior reflects what each individual agent in fact believes and the epistemic advantages of doing so outweigh the pragmatic disadvantages of doing so.

At this point, someone might worry that one of our key examples—the Classroom Case—does not support the verdict that it is rational all things considered for the relevant agents to conform to (iii). For, in that case, the epistemic disadvantages of conforming to (iii) *might* outweigh the pragmatic advantages of doing so. Or, to put it more specifically, while it may be pragmatically advantageous to stay quiet and steer clear of being identified as the only person who has not understood the material, it

may nonetheless be very epistemically advantageous to speak up. In particular, one might learn a lot by doing so.

Let us make three comments concerning these issues. First, it is correct that some cases may be construed in a way that ensures that the epistemic disadvantages of conforming to (iii) outweigh the pragmatic advantages of doing so. But, secondly, it is not clear that the Classroom Case is this kind of case. As formulated above, the Classroom Case is not described in sufficient detail to deliver a firm verdict on what the epistemic consequences of staying quiet are. This would depend on the further features of the situation—for example, the subject matter being taught and who is teaching it. If the subject matter is such that knowing it would result in a large body of new knowledge, one would miss out on a lot of knowledge by staying quiet. On the other hand, there are also cases in which one would not, epistemically speaking, be missing out. The subject matter might be so esoteric or isolated that one would not get any noteworthy epistemic gain from learning about it. Third, even if there are cases in which the epistemic disadvantages of conforming to (iii) outweigh the pragmatic advantages of doing so, our case is not undermined. For we must show only that there need not be any error of rationality, all things considered, involved with respect to (iii) in a case where the pragmatic payoff associated with conforming to (iii) exceeds the epistemic disadvantage of doing so. But here it is easy to see that there are many more fully specified versions of the Classroom Case that would do—just consider a case in which the subject matter is highly esoteric, or in which the teacher will not be helpful or informative at all if he finds out that someone did not understand the material. If so, there is no strong reason to doubt that pluralistic ignorance can be a rational phenomenon even in cases like the Classroom Case.

Thus we have argued that no errors of rationality, all things considered, need be involved in situations of pluralistic ignorance. This is an interesting conclusion. Among other things, if pluralistic ignorance can arise in groups of rational agents, there are no obvious reasons why philosophical tools from logic and game theory should not be able to help us model and reason about the phenomenon. We will indicate how below.

5 Concluding remarks

In this paper, we have addressed several issues that we take to be important prerequisites for a formal, social epistemological analysis of pluralistic ignorance. First, we disentangled several characterizations of pluralistic ignorance and provided our reasons for preferring (PI_4). In particular, in contrast to other common characterizations of pluralistic ignorance, we argued that (PI_4) manages to reflect the *social epistemological* nature of the phenomenon.

Second, based on (PI_4), we argued that pluralistic ignorance can arise in groups of perfectly rational agents. Although there are ways rational agents may reason to dissolve situations of pluralistic ignorance, we showed that it can be rational for agents, all things considered, to dismiss this reasoning. In particular, when the two necessary conditions (Con 1) and (Con 2) are met, pluralistic ignorance can arise and persist in groups of rational agents. In contrast to the socio-psychological literature that seems to assume that people in situations of pluralistic ignorance make *errors* of various

sorts, our arguments show that they need not make any errors of rationality. As such, we can in many cases view pluralistic ignorance as a side effect of certain features of the *epistemic* and *social interactive* context rather than as a side effect of certain psychological, sociological, and cognitive errors.

Throughout the paper we have assumed that all agents in situations of pluralistic ignorance interact with and observe each other's behavior. But it has been documented that pluralistic ignorance can arise in more complex social network structures, where only some agents in the structure interact with each other. Centola et al. (2005), for instance, use computational agent-based models to investigate how unpopular social norms can emerge in various social network structures. They investigate networks in which all people need not be connected—need not be able to observe and interact with each other—and in which people differ in their convictions of the desirability of the relevant social norm. They show that if a minority of people in a group endorse an otherwise highly unpopular social norm with conviction, and if the social network consists of adjacent neighborhoods, then this endorsement can quickly spread from the local group to the whole social network. In such cases, pluralistic ignorance can arise in the social network even when some agents are not directly interacting and observing each other.[16] To fully understand the epistemic nature of pluralistic ignorance—and in particular, to understand how pluralistic ignorance emerges—such internal mechanisms in complex social groups need further investigation. More generally, there is much epistemological work to be done in analyzing how epistemic concepts such as *trust*, *testimony*, and *informational cascades* might influence the emergence of pluralistic ignorance.

5.1 Ramifications for modeling and dissolving situations of pluralistic ignorance

Our paper has ramifications both for how to formally model pluralistic ignorance and for how to dissolve situations of pluralistic ignorance. We go through each in turn.

To capture the central epistemic components of pluralistic ignorance, a formal framework need be able to accommodate how beliefs can be formed based on observations of actions in a social context. *Dynamic epistemic logic* seems like a good candidate for such a framework. In dynamic epistemic logic, one can model how agents' beliefs—including higher-order beliefs—may change as a result of certain actions such as observations and public announcements. Although a full discussion of the prospects of using dynamic epistemic logic to model pluralistic ignorance is beyond the scope of this paper, we should note that simple versions of dynamic epistemic logic have already been used to model aspects of the phenomenon. Hendricks (2010), for instance, makes a first step towards modeling pluralistic ignorance using a combination of *public announcement logic* and *formal learning theory*.[17] In the paper, he aims to show that if certain types of knowledge transmissibility are possible, then (his version of) pluralistic ignorance can be dissolved by public announcements. It

[16] Interestingly, however, such cascades need not arise in social networks where everyone is connected.

[17] Public announcement logics are simple versions of dynamic epistemic logic where only one type of epistemic action, namely public announcement, is considered.

should be noted, though, that Hendrick's view on pluralistic ignorance seems to differ somewhat from the standard views mentioned in Sect. 3.

Hansen (2012) also investigates how pluralistic ignorance can be dissolved using a version of public announcement logic combined with beliefs. He shows that a type (PI_2) definition of pluralistic ignorance—in contrast to the prevailing view in the literature—can be a quite stable phenomenon in the sense that even if some members of the group publicly announce their private beliefs, pluralistic ignorance need not thereby be dissolved. To the knowledge of the authors, however, no one has yet attempted to model how pluralistic ignorance can *arise* using logic and game theory. So there is plenty of logical work to be done in the future.

Our discussion of pluralistic ignorance also serves to point out certain missing features in available models of pluralistic ignorance. Consider, for instance, the formal models of pluralistic ignorance in Centola et al. (2005) and Lisciandra et al. (2011). Centola et al. (2005), as mentioned, provide a formal model of how pluralistic ignorance can arise as a product of minority opinions being propagated through informational cascades in certain types of social networks. In their model, however, we cannot make a distinction between what an agent believes about other agents in the social group and how these others agents publicly act. That is, what an agent believes about other agents is exhausted by how the other agents act. This means that agents' in the model can never come to learn—nor doubt—that other agents are acting contrary to their private beliefs. But given our discussion in Sect. 4, there are cases of pluralistic ignorance in which agents plausibly come to believe that other agents in the group act contrary to what they actually believe. Centola et al. (2005) cannot model such cases.

Lisciandra et al. (2011) are aiming to develop a formal model of how pluralistic ignorance can arise in groups of rational Bayesian agents. Like us, they stress that belief formations based on observations of public behavior play an important role in situations of pluralistic ignorance. Unlike us, but much like Centola et al. (2005), however, they seem to assume that an agent's beliefs (credences) about other agents' private beliefs are exhausted by observing how they behave. In case their model works—the verdict on whether it does is still open—they still need to justify that belief formation based on observations is always a *rational* process in situations of pluralistic ignorance. Yet, to do so, they need to explain how agents can dismiss the reasoning step (d) in the (PI-Rat) argument. But since it is *not* possible for agents in their model to come to believe that other agents might not act in accordance with their private beliefs, they cannot *explain* how agents can entertain the reasoning in (d) and yet dismiss it by epistemic means. For it is hardwired into their model that agents cannot entertain the relevant critical reasoning in (d). As such, their model—if ultimately successful—will only show that pluralistic ignorance can arise among rational Bayesian agents, *if* one makes the further assumption that it is always rational for agents to take the actions of others as truly reflecting what they believe. But in Sect. 4 we gave reasons for why we should *not* make this assumption in all situations of pluralistic ignorance. If so, there are reasons to hold that Lisciandra et al. (2011) should construct their model of pluralistic ignorance without the assumption. As it currently stands, the model does not satisfy this desideratum.

In addition to help shaping what a formal model of pluralistic ignorance should look like, our paper also shows how a precise understanding of pluralistic ignorance

can help explain how the phenomenon might be dissolved. Given the many potentially harmful side effects of pluralistic ignorance, considerable efforts have been put into investigations of how it can be dissolved. Many have proposed that pluralistic ignorance can be dissolved simply by making people in the relevant social context aware of the possibility that they are in a situation of pluralistic ignorance. For instance, Schroeder and Prentice (1998) show that college students can be made to drink considerably less if they are informed about pluralistic ignorance. As a consequence of our discussion in Sect. 4, however, it is clear that merely informing people in a situation of pluralistic ignorance about the possibility need not dissolve it. On the one hand, if the people involved in the situation are rational, they will already be aware of the possibility of pluralistic ignorance—in form of the reasoning step (d) in (PI-Rat). To dissolve pluralistic ignorance in groups of such agents, they need enough additional evidence to make it epistemically rational for them to assign a high credence to the discrepancy in (d) between private belief and public behavior. On the other hand, if the people involved in the situation are not rational, they need not be aware of the reasoning in (d), and informing them about the possibility of pluralistic ignorance is no guarantee that they will thereby be able to dissolve it by rational reflection. In groups of such agents, one needs first "teach" them the reasoning in (d), and then proceed as in the case of groups of rational agents.

So there is much more work to be done in understanding the epistemic and social interactive nature of pluralistic ignorance. Hopefully, our paper has helped shape some of the directions that this research could and should take.

Acknowledgments First of all, we would like to thank Carlo Proietti and Frank Zenker for organizing the CPH-LU workshops on social epistemology. We would also like to thank the participants of the CPH-LU workshops in Lund and Copenhagen and the participants of the LIRa/LogiCIC seminar at ILLC for useful comments and suggestions. We benefited especially from discussions on the topic with Carlo Proietti, Rasmus Rendsvig, and Vincent Hendricks. Finally, we would like to thank two anonymous reviewers for useful comments and suggestions.Jens Ulrik Hansen is sponsored by the Carlsberg Foundation.

References

Bicchieri, C. (2006). *The grammar of society—The nature and dynamics of social norms*. Cambridge: Cambridge University Press.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, *100*(5), 992–1026.

Borsari, B., & Carey, K. B. (2001). Peer influences on college drinking: A review of the research. *Journal of Substance Abuse*, *13*, 391–424.

Centola, D., Willer, R., & Macy, M. (2005). The emperor's dilemma. A computational model of self-enforcing norms. *American Journal of Sociology*, *110*(4), 1009–1040.

Christensen, D. (2004). *Putting logic in its place—Formal constraints on rational belief*. Oxford: Clarendon Press.

Danmarks Statistik. (2011). Befolkningens brug af internet—2010. http://www.dst.dk/pukora/epub/upload/15239/it.pdf.

Halbesleben, J. R. B., & Buckley, M. R. (2004). Pluralistic ignorance: Historical development and organizational applications. *Management Decision*, *42*(1), 126–138.

Halbesleben, J. R. B., Wheeler, A. R., & Buckley, M. R. (2007). Understanding pluralistic ignorance: Application and theory. *Journal of Managerial Psychology*, *22*(1), 65–83.

Hansen, J. U. (2012). A logic-based approach to pluralistic ignorance. In J. D. Vuyst & L. Demey (Ed.), Future directions for logic—Proceedings of PhDs in Logic III: College Publications, pp. 67–80.

Hendricks, V. F. (2010). Knowledge transmissibility and pluralistic ignorance: A first stab. *Metaphilosophy*, *41*(3), 279–291.

Hendricks, V. F., Olsson, E. J., Zenker, F. (2012). Copenhagen-Lund workshop in social epistemology, 2010–2011. http://www.fil.lu.se/conferences/conference.asp?id=43&lang=se.

Katz, D., & Allport, F. H. (1931). *Student attitudes*. Syracuse, NY: The Craftsman Press.

Krech, D., & Crutchfield, R. S. (1948). *Theories and problems of social psychology*. New York: McGraw-Hill.

Lisciandra, C., Muldoon, R., Hartmann, S., Rad, S. R. (2011). Pluralistic ignorance, slides from the 4th copenhagen lund workshop in social epistemology. http://www.fil.lu.se/files/conference214.pdf.

Miller, D. T., & McFarland, C. (1987). Pluralistic ignorance; when similarity is interpreted as dissimilarity. *Journal of Personality and Social Psychology*, *53*, 298–305.

Miller, D. T., & McFarland, C. (1991). When social comparison goes awry: The case of pluralistic ignorance. In J. Suls & T. Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 287–313). Hillsdale, NJ: Erlbaum.

O'Gorman, H. J. (1986). The discovery of pluralistic ignorance: An ironic lesson. *Journal of the History of the Behavioral Sciences*, *22*(October), 333–347.

Prentice, D. A., & Miller, D. T. (1993). Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology*, *64*(2), 243–256.

Schroeder, C. M., & Prentice, D. A. (1998). Exposing pluralistic ignorance to reduce alcohol use among college students. *Journal of Applied Social Psychology*, *28*(23), 2150–2180.

Westphal, J. D., & Bednar, M. K. (2005). Pluralistic ignorance in corporate boards and firms' strategic persistence in response to low firm performance. *Administrative Science Quarterly*, *50*(2), 262–298.