Intersubjective belief[*]

Vivienne Brown

Open University, UK

(v.w.brown@open.ac.uk)

Abstract

This paper proposes a new model of shared belief amongst individual subjects based on a new approach to theorising individual subjects in social context. In this approach, which I term the *intersubjective* approach, individual subjects are modelled in terms of the standpoint of *each of us*, thereby incorporating the phenomenological standpoint of an individual subject's inclusion of herself within the plurality, 'us' (a class in the distributive sense). This provides resources for a new model of shared belief, including common belief, in terms of *intersubjective belief*, which is an individual subject's belief that 'each of us has the same belief that *p*'. The paper argues that the intersubjective model of shared belief provides a non-reductive alternative to the standard interactive model of mutual belief and common belief, and so provides a non-individualistic framework for analysing shared belief in social contexts. As an illustration, the intersubjective model of common belief is applied to the Hi-Lo game; the solution is (*High*, *High*).

I Shared belief

As social beings we take it for granted that we inhabit a shared world. Unless a person lives as a total social recluse, it is a characteristic feature of their life that it includes shared situations, shared beliefs, shared discussions, and so forth, with others. This raises a fundamental question of how to model individual subjects' cognizance of their shared world. This paper proposes a new approach to this question and uses it to develop a new model of shared belief, including common belief.

The simplest kind of shared belief is that every subject in some population believes that *p*, but this does not take into account subjects' beliefs about their shared beliefs. The question of how to model subjects' beliefs about their shared beliefs goes to the heart of the question of how to model subjects' cognizance of their shared world. The standard approach, sometimes termed the 'interactive' approach, models these beliefs in terms of hierarchies of individual beliefs. According to the interactive approach, mutual belief amongst subjects in some population is defined as a finite hierarchy of beliefs, such that, expressed informally, first-order mutual belief is 'every subject believes that *p*', second-order mutual belief is 'every subject believes that *p*, and every subject believes that every subject believes that *p*', and so on; and common belief is defined as an infinite hierarchy (or recursive system) of beliefs, such that 'every subject believes that *p*, every subject believes that every subject believes that *p*, … and so on, *ad infinitum*' (equivalently, 'every subject believes "*p* and there is common belief that *p*"'). This implies a reductive view of shared belief because mutual and common belief that *p* are explained solely in terms of individual beliefs that *p* (to some specified order) such as: I believe '*p* and you believe that *p* and you believe that I believe that *p*'. These interactive notions of mutual and common belief, and formal specifications of them, have been adopted as standard across academic disciplines including philosophy, the social sciences, game theory, computer science and artificial intelligence.[1]

This paper develops a new approach to theorising individual subjects' cognizance of their shared world and it argues that this provides resources for a new model of shared belief in which individual beliefs that *p* (of whatever order) are not sufficient for shared belief. Individual subjects in a shared social context are theorised in terms of the standpoint of *each of us*, not in terms of 'I' and 'you'. As 'each' and 'each of us' are grammatically singular, subjects are theorised as singular entities, but each subject includes herself within 'us'. I term this the *intersubjective* approach. By displacing the distinction between 'I' and 'you', the intersubjective approach provides a new framework for analysing individual subjects' understanding of what is shared amongst 'us'. This in turn provides resources for a new model of shared belief by displacing the distinction between 'I believe' and 'you believe' that is characteristic of the interactive approach yet without relying on a collectivist interpretation of 'we believe'. The paper argues that the intersubjective approach provides a non-reductive

---

[1] Lewis (2002/1969) and Aumann (1976, 1999) are canonical. Weirich (2011) provides a brief introduction to 'interactive epistemology', and Vanderschraaf and Sillari (2014) provide a survey of common knowledge.

analysis of shared belief amongst individual subjects that is simpler and more plausible than the interactive approach and which yields fresh resources for application to social issues.

Section II introduces the intersubjective approach of *each of us*. Section III introduces the model of *intersubjective belief*. Sections IV and V extend the model of intersubjective belief, first to a general model of shared belief (*plural intersubjective belief*) and then to common belief (*intersubjective common belief*). Section VI concludes the paper with an application of the intersubjective model to a simple coordination problem, the Hi-Lo game.

II Introducing the intersubjective approach: *each of us*

A challenge in understanding individual human subjects in shared social contexts is how to theorise the relation between individual subjects and other subjects. Theorising an individual subject in terms of 'I' seems to set up a distance between that subject and other subjects, a distance that then has to be bridged in some way. For example, if the relation is construed in terms of 'I' and 'you', this seems to set up a distance between 'I' and 'you' that then has to be bridged. The intersubjective approach proposed in this paper responds to this challenge by theorising an individual subject's relation with others in terms of the standpoint of *each of us*.[2] An advantage of this formulation is that it registers the way that a subject includes herself equivalently with each of the other subjects in 'us', so there is no distance set up between an individual subject and other subjects that then has to be bridged.

The intersubjective approach theorises individual subjects as singular entities. Grammatically 'each' is a distributive pronoun that individuates 'us', so that 'each' and 'each of us' are grammatically singular. Yet as 'each of us' includes the subject herself as one of 'us', individual subjects are theorised as from their own standpoint of self-inclusion in 'us'. The expression 'each of us' thus registers the phenomenology of the reflexive standpoint for singular subjects who see themselves as one of 'us'.

As the object pronoun 'us' is indexical, its reference is context-sensitive.[3] The linguistic meaning of 'us', which is invariant to context, is 'a plurality of subjects that includes me', where 'plurality' here is a class in the distributive sense. The referent of 'us' in a context is a

---

[2] I take it that *each of us* is expressible in languages other than English. The argument of this paper is not meant to be dependent on the particularities of the English language or on vernacular usage of English.
[3] The argument of this paper is meant to be independent of any particular philosophical understanding of indexicality. Kaplan (1989) is an influential study of indexicality; Braun (2016) provides a survey.

class or reference group (in the distributive sense) within which the subject includes herself and others. The expression 'each of us' is also indexical. Its linguistic meaning, also invariant to context, is 'each of the plurality, "us", which plurality includes me'. The referents of 'each of us' in a context are the members of the class or reference group that is the referent of 'us', amongst whom the subject includes herself equivalently with each of the other members. For 'each of us' in a specific context, 'each of us' should be read as 'each of us (in R)' or 'each of us (the members of R)', where 'R' is the specific reference group whose members, amongst whom the subject includes herself, comprise the referents of 'each of us'. This implies that the subject of 'each of us (in R)' is cognizant[4] of the criteria that define this reference group, R, within which she includes herself, otherwise she could not know which subjects to include; that is, she could not know who are the referents of 'each of us'. For example, in 'each of us (in reference group, "the Philosophy Department at Interstate University") is a philosopher', the subject is cognizant that membership of the Philosophy Department at Interstate University is the criterion for inclusion in the reference group, and that the individual members of the Philosophy Department at Interstate University, amongst whom she includes herself, are the referents of 'each of us'.

In taking herself to be a member of the inclusive plurality, 'us', a subject thus includes herself equivalently with each of the others in 'us'. This implies that each subject in the reference group is taken by the subject to have the requisite characteristics of subjecthood in order to qualify for inclusion in 'us'. Furthermore, anything the subject holds true of 'each of us', she holds true of herself. I term this the *Principle of Intersubjective Inclusion*, viz. that anything a subject holds true of 'each of us', she holds true of herself because she includes herself in 'us'. In the example just cited, if a subject believes 'each of us (in reference group, "the Philosophy Department at Interstate University") is a philosopher', she believes it of herself that she is a member of the Philosophy Department at Interstate University and is a philosopher.[5]

As this example suggests, the reference group, R, might in practice take any one of a number of forms, including a loose assemblage of subjects (e.g. residents of a town), an informal

---

[4] I abstract from issues as to whether this cognizance is a form of knowledge, belief, understanding, awareness, and so forth.

[5] Although 'each of us' is inclusive with respect to those who are the members of 'us', it is exclusive with respect to those who are deemed not to be one of 'us'. For example, if 'each of us (in reference group, "the Philosophy Department at Interstate University") is a philosopher but Pierre is a physicist', Pierre is excluded from being one of 'us'.

grouping (e.g. people at a party), a formal group (e.g. a committee), or a discourse community (e.g. the community of academic physicists). Whatever the features of R, individual subjects in a specific context are theorised as situating themselves within that particular reference group.[6]

Theorising individual subjects in terms of 'each of us', rather than 'I' and 'you', facilitates an analysis of individual subjects as singular subjects and as members of a class (in the distributive sense) within which they include themselves. The intersubjective approach thus analyses individual subjects from their own reflexive standpoint of self-inclusion in the plurality, 'us', and so provides a means of analysing subjects' self-understanding in a world they recognise to be shared with others.

'Each of us' is not equivalent to 'every' or 'each' of universal quantification, according to which, if set $A = \{a_1, a_2, \dots , a_n\}$, 'every member of A is F' is equivalent to 'a_1 is F & a_2 is F & … & a_n is F'. Here items are sorted or collected into a class or set by a theorist or external observer. The members of A do not have cognizance of set A itself as this is knowledge that belongs to the theorist or external observer who has collected them into a set as objects of analysis or observation. For example, in the case of 'every ball in box B is red', no individual ball in box B cognizes that it has itself been collected into the set of balls in box B or that every ball in box B is red. Similarly, in the case of 'every person in park P is happy', no individual person in park P cognizes that she has been collected into the set of persons in park P or that every person in park P is happy. In this latter case, although the individual entities might be thought to have the cognitive capacities to understand such quantification, the model does not incorporate such understanding for members of the set. According to the model, persons in park P have no more cognizance of the quantification than balls in box B have.

In contrast, 'each of us (in park P) is happy' implies a subject who is cognizant that 'each of us (in park P)' is classed as happy. This involves a kind of universal quantification but with the difference that the quantification is made by one of the entities being quantified over, so it is no longer the case that none of the entities being quantified over has cognizance of the quantification. To differentiate between these two kinds of universal quantification, I term the former 'external universal quantification' and the latter 'internal universal quantification'. External universal quantification applies where the quantification is made by a theorist or observer who is external to the set of entities being quantified over; 'every ball in box B is red'

---

[6] This paper does not consider the issue of group belief where 'group' is conceived in a strictly collective or holistic sense, for example, a government, NGO, or corporation.

and 'every person in park P is happy' are examples of external universal quantification. Internal universal quantification applies where the quantification is made by a member (or members) of the set being quantified over; it includes self-quantification and implies the subject's (or subjects') cognizance of the set itself.

The intersubjective standpoint involves internal universal quantification and implies a subject's cognizance of the group, 'us', within which she includes herself. This phenomenology of self-inclusion comprises an essential element of the subject's self-understanding of her relation with others in a shared context. This enables the intersubjective approach to provide an analysis of the inclusive plurality, 'us', without introducing issues of collectivism or holism. This in turn provides an interpretation of the distributive 'we' in ordinary English: 'we are F' is analysed as 'each of us (in R) is F'; for example, 'we are happy' is analysed as 'each of us (in R) is happy'.

It will be argued in the course of this paper that the intersubjective approach provides resources for a new non-reductive model of shared belief amongst individual subjects.

III Intersubjective belief

Shared belief is theorised in the intersubjective approach in terms of subjects' beliefs that the belief that p is the same for 'each of us'. This integrates the singularity of individual subjects who hold the belief and the inclusive plurality, 'us', within which it is believed that the belief that p is shared.

A key notion is *intersubjective belief*, which is a single subject's belief that 'each of us has the same belief that p'.[7] In a specific context, the intersubjective belief that p held by a subject who includes herself as a member of reference group, R, is her belief that 'each of us (in reference group, R) has the same belief that p':

> In a specific context, subject i, who includes herself in reference group, R, holds the **intersubjective belief** that p iff i believes 'each of us (in reference group, R) has the same belief that p'.

For example, subject i who includes herself in R (composed of those who believe that philosophy is interesting) holds the intersubjective belief that philosophy is interesting iff i

---

[7] The definition of intersubjective belief is meant to be independent of any particular philosophical understanding of belief.

believes 'each of us (in R) has the same belief that philosophy is interesting'. A philosopher in the philosophy department at Interstate University holds the intersubjective belief that philosophy is going through a period of change iff she believes 'each of us (in the philosophy department at Interstate University) has the same belief that philosophy is going through a period of change'.

Adapting standard notation, the intersubjective belief that p of subject i who includes herself in reference group, R, is given in (1) as $IB_{i/R}(p)$. $B_{i/R}$ is the belief operator and $\bar{B}p$ is 'each of us (in R) has the same belief that p':[8]

$$IB_{i/R}(p) \equiv B_{i/R}(\bar{B}p). \tag{1}$$

This reads: the intersubjective belief that p of subject i who includes herself in reference group, R, is i's belief that 'each of us (in R) has the same belief that p'.[9] The content of the subject's belief, viz. 'each of us (in R) has the same belief that p', given as $\bar{B}p$ in (1), provides an example of internal universal quantification because the quantification over the members of set R is performed by a member of R.[10]

By the Principle of Intersubjective Inclusion, if a subject holds the intersubjective belief (hereafter IB) that p, then the subject believes it of herself that she believes that p. In other words, as the subject includes herself when quantifying over the members of R, anything she attributes to the members of R she attributes to herself. For example, if subject i believes 'each of us (in R) has the same belief that philosophy is interesting', then subject i believes it of herself that she believes that philosophy is interesting.[11]

---

[8] In a particular instance where R comprises just one subject, an intersubjective belief that p reduces to a subject's belief that only she herself believes that p, viz. $B_i(Bp)$, where 'B' refers reflexively to the belief of the subject herself.

[9] Partially shared beliefs involve a weakening of 'each of us', for example, 'nearly each of us' or 'hardly any one of us'. The corresponding weakened intersubjective beliefs would be the belief that 'nearly each of us (in R) has the same belief that p' or 'hardly any one of us (in R) has the same belief that p'.

[10] Although this paper focuses on intersubjective *belief*, an analogous analysis can be made of intersubjective *knowledge* as a subject's knowledge that 'each of us (in R) has the same knowledge that p', viz. $IK_{i/R}(p) \equiv K_{i/R}(\bar{K}p)$ (cf. n. 23). Analogously with intersubjective belief, this definition of intersubjective knowledge is meant to be independent of any particular philosophical understanding of knowledge.

[11] If a subject in R believes 'not one of us (in R) believes that p', she does not believe that p. This is a null IB, viz. $IB^-_{i/R}(p) \equiv B_{i/R}(\neg\bar{B}p)$. If a subject in R believes 'each of us (in R) has the same belief that not-p', this is given as $IB_{i/R}(\neg p) \equiv B_{i/R}(\bar{B}\neg p)$.

For subjects in R to have the *same* belief that p requires that both the linguistic meaning and content of the linguistic expression for p are invariant across the subjects in the given R. This suggests that indexicals such as 'I' and 'you' do not qualify for inclusion. For example, if 'p' is 'I am hungry' or 'you are working late', the content of these expressions depends on which subject holds the belief. In the case of 'each of us (in R) has the same belief "I am hungry"', the content of the belief is different for each member of R: i believes 'each of us has the same belief that i is hungry, j believes 'each of us has the same belief that j is hungry', and so on. If just one subject in R is hungry, there could be an IB that 'one of us (in R), namely subject k, is hungry', as here the content of 'p' would be invariant across subjects in a given R. Once R is given, however, 'p' can include 'each of us' because the referents of 'each of us' are invariant across subjects in a given R. For example, the belief that 'each of us (in R) is hungry' could be held by any one of the subjects in R, with the same content as well as the same linguistic meaning. Importantly, an IB may thus be about 'each of us', as in the IB that 'each of us (in R) is hungry'.[12]

As an IB that p is the belief that 'each of us (in R) has the same belief that p', such a belief is true iff each subject in R has the same belief that p. Thus the truth of an IB is independent of whether p is true: an IB might be true even if p is false (each subject has the same false belief that p), or it may be false even if p is true (it is not the case that each subject has the same true belief that p). In other words, the truth of an IB depends on whether the belief that p is shared amongst those in R, and this is independent of whether p is true. This is illustrated by the notation for a true IB that p, which is $IB_{i/R}^{T}(p) \equiv B_{i/R}^{T}(\bar{B}p)$, where the superscript 'T' denotes the truth of the belief. The truth of i's IB that p depends on the truth of i's belief that 'each of us (in R) has the same belief that p'. It does not depend on the truth of p. If the belief that p were to be marked as true this would be written as '$\bar{B}^{T}p$', which is 'each of us (in R) has the same true belief that p'. In this paper it is left open as to whether p is true, so '$\bar{B}p$' is not written with a 'T' superscript.

As noted above, to include herself equivalently with others in the inclusive plurality, 'us', a subject takes it that each of the subjects in 'us' has the requisite characteristics of subjecthood, otherwise she could not include them all equivalently. Similarly, to hold the IB that p, a subject

---

[12] The intersubjective approach is not restricted to understanding shared beliefs, as recognition of difference in beliefs amongst 'us' is also possible, for example, the IB that 'each of us (in R) has a different belief about event E' is the belief that 'each of us (in R) has the same belief that each of us (in R) has a different belief about event E'.

regards herself and the other subjects in R equivalently as doxastic subjects who are each capable of holding the same belief that p, otherwise she could not believe that 'each of us' has the same belief that p. To hold the IB that p in a context thus implies that the subject believes that 'each of us (in R)' has the appropriate cognitive capacities, information, and so forth, to hold the same belief that p. For example, the belief that 'each of us (parents at school S) has the same belief that the teachers at S are wonderful' requires different cognitive capacities and information, and so forth, of subjects compared with the belief that 'each of us (academic physicists) has the same belief that cosmology is interesting'. As empirical subjects are members of different doxastic communities, $R_1, R_2, \dots , R_n$, the IBs relating to these different communities might require different assumptions about the cognitive capacities, information, and so forth, of doxastic subjects included in R.

IBs may be taken to a higher degree of intersubjectivity, where the degree of intersubjectivity is given by the number of occurrences of 'each of us (in R) has the same belief that'. The IBs discussed so far are of first-degree intersubjectivity (they are first-degree IBs). In the case of a second-degree IB, a subject believes 'each of us (in R) has the same belief that each of us (in R) has the same belief that p', which is equivalent to the belief that 'each of us (in R) has the same IB that p'. Second-degree IB of subject i in R is given as $IB^2_{i/R}(p) \equiv B_{i/R}(\bar{B}^2 p) \equiv B_{i/R}(\overline{IB}p)$, where $B_{i/R}(\bar{B}^2 p)$ is 'i believes "each of us (in R) has the same belief that each of us (in R) has the same belief that p"', and $B_{i/R}(\overline{IB}p)$ is 'i believes "each of us (in R) has the same IB that p"'. A second-degree IB that p is true iff each subject (in R) has the same true first-degree IB that p.

 An IB of dth-degree intersubjectivity for subject i who includes herself in R is given in (2) as $IB^d_{i/R}(p)$. In the context of human beliefs d is restricted to finite values:

$$IB^d_{i/R}(p) \equiv B_{i/R}(\bar{B}^d p) \equiv B_{i/R}(\overline{IB}^{d-1}p) \qquad 1 \leq d < \infty . \qquad (2)^{[13]}$$

A dth-degree IB that p is true iff each subject in R has the same true $(d-1)^{\text{st}}$-degree IB that p. In practice for human subjects IBs are likely to be of low degree of intersubjectivity. The argument of this paper requires only second-degree IBs.

---

[13] By definition, $IB^1_{i/R}(p) \equiv B_{i/R}(\bar{B}p) \equiv B_{i/R}(\overline{IB}^0 p)$.

It is to be noted that an IB does not logically require an IB of higher degree; that is, $\bar{B}^d p$ does not logically imply $\bar{B}^{d+1} p$. For example, the leader of a research team might be responsible for conveying information to the other members of the research team, who do not communicate with one other, so that if the team leader informs each of them individually that p, she believes 'each of us (members of the research team) has the same belief that p', but she does not hold a second-degree IB because that would require each of the other members of the team to hold an IB that p (not just the belief that p). As the team leader has told each of them individually only that p, and as they do not communicate with each other, they do not believe 'each of us (members of the research team) has the same belief that p'. The team leader therefore does not have second-degree IB that p. Her first-degree IB that p does not logically require a second-degree IB that p.

IV Plural intersubjective belief

I have so far considered the IB of a single subject in R and I have argued that the content of the belief that 'each of us (in R) has the same belief that p' involves internal universal quantification by the subject who holds the IB. If every subject in R holds the same true IB that p, there is *plural intersubjective belief* amongst subjects in R. The notion of plural intersubjective belief involves external universal quantification over subjects in R (every subject in R holds the same IB that p)[14] as well as internal universal quantification by subjects who hold the IB ('each of us (in R) has the same belief that p').

Plural intersubjective belief that p amongst subjects in R is given in (3) as $PIB_R(p)$, where $\overline{IB}_R^T(p)$ is 'every subject in R [i.e. external universal quantification] holds the same true IB that p' and $\bar{B}_R^T(\bar{B}p)$ is 'every subject in R [i.e. external universal quantification] has the same true belief that "each of us (in R) [i.e. internal universal quantification] has the same belief that p"'.[15] The superscript 'T' denotes the truth of the belief:

$$PIB_R(p) \equiv \overline{IB}_R^T(p) \equiv \bar{B}_R^T(\bar{B}p) \tag{3}$$

Plural intersubjective belief (hereafter plural IB) is possible with true or false p, just as IBs are possible with true or false p. In (3) the shared IB is given as true so that the belief which every

---

[14] Plural IB could be weakened to 'most' or 'some' of the members in R hold the same IB.

[15] The notational convention is that 'every subject in R has the same belief that' is given by $\bar{B}_R(\ )$, whereas 'each of us (in R) has the same belief [or same IB] that p', which is always in the content of a belief, is given by $(\bar{B}p)$ or $(\overline{IB}p)$.

subject has, that 'each of us (in R) has the same belief that p', is also given as true; but p is not given as true. Although the belief that 'each of us (in R) has the same false belief that p' is ruled out (it is assumed that a subject does not believe herself to hold a false belief), it might be that subjects are mistaken in believing that p even though they truly believe that they each have the same belief that p. Thus plural IB is possible with p true or false; for example, there might be plural IB that the proposed meeting is to take place in room AA (p is true) or that the proposed meeting is to take place in room BB (p is false). The former is a true plural IB and the latter is a false plural IB.

Although an individual subject might be mistaken in believing 'each of us (in R) has the same belief that p' (irrespective of whether p is true of false), if every subject in R has the same IB that p, then every subject in R believes that p so their IBs are true. In other words it can't be the case that every subject holds the same IB that p (and is not mistaken about her own belief) yet those IBs are not true. (There is thus an element of tautology in the explanation of plural IB given above.[16]) Conversely, if every subject in R believes 'not one of us (in R) believes that p', then no-one in R believes that p; this is 'null plural IB'. [17,18]

As with IBs, plural IBs can be taken to any degree of intersubjectivity:

> There is d^{th}-degree **plural intersubjective belief** that p amongst subjects who include themselves in reference group, R, iff every subject in R holds the same true d^{th}-degree IB that p.

Plural IB that p of dth-degree amongst subjects who include themselves in R is given in (4) as $PIB_R^d(p)$, where $\overline{IB}_R^{d,T}(p)$ is 'every subject in R has the same true d^{th}-degree IB that p' and $\bar{B}_R^T(\overline{IB}^{d-1,T}p)$ is 'every subject in R has the same true belief that "each of us (in R) has the same true $(d-1)^{\text{st}}$ IB that p"'. As with IBs, d is restricted to finite values for human beliefs:

$$PIB_R^d(p) \equiv \overline{IB}_R^{d,T}(p) \equiv \bar{B}_R^T(\overline{IB}^{d-1,T}p) \qquad (1 \leq d < \infty). \qquad (4)$$

---

[16]The tautology would disappear if IBs were weakened as suggested in notes 9, 14.

[17] Null plural IB is $PIB_R^-(p) \equiv \bar{B}_R(\neg\bar{B}p)$, in contrast with plural IB that not-p', viz. $PIB_R(\neg p) \equiv \bar{B}_R(\bar{B}\neg p)$. Cf. note 11.

[18] If every subject believes 'each of the *other* subjects (in R) believes that p', and does not herself believe that p, then no one believes that p, and the IBs are false. This lack of awareness of others' beliefs is most likely in repressive societies where subjects are not aware of others' actual beliefs but rely on 'official' views as to what people believe.

In practice only low degrees of intersubjectivity are likely to be relevant for human beliefs. The argument of this paper requires only second-degree plural IB, viz. $PIB_R^2(p) \equiv \overline{IB}_R^{2,T}(p) \equiv \overline{B}_R^T(\overline{IB}^T p)$, just as it requires only second-degree IBs.

As with IBs, plural IB does not logically require plural IB of a higher degree. For example, returning to the leader of the research team who is responsible for conveying information to the other members of the team who do not communicate with each other, the team leader this time informs each of the other members individually not only that p but also that she is informing each of them individually that p. Now every member of the team has first-degree IB that p (every team member has the same belief that 'each of us (members of the research team) has the same belief that p'). There is thus first-degree plural IB that p amongst the members of the team. But team members do not have second-degree IB that p; they do not all believe 'each of us (members of the research team) has the same belief that each of us (members of the research team) has the same belief that p'. The team leader alone has second-degree IB that p so there is no second-degree plural IB that p amongst the members of the research team.

Plural IB provides a new model of shared belief according to which subjects not only *have the same beliefs* (irrespective of the degree of belief) but also *believe that they each have the same beliefs* (irrespective of the degree of belief). Plural IB thus has the following characteristics of shared belief (*same* beliefs as above, that is, linguistic meaning and content of 'p' are invariant across subjects in a given R):

Characteristics of shared belief that p

    (i)   Every subject (in R) has the same belief that p.

    (ii)   Every subject (in R) has the same true belief that they each have the same belief that p.

    (iii)   (ii) holds irrespective of the degree of belief.

In the case of first-degree plural IB that p, (i) every subject (in R) has the same belief that p, and (ii) every subject (in R) has the same true first-degree IB that p. In the case of second-degree plural IB that p, (i) every subject (in R) has the same belief that p, and (ii) every subject (in R) has the same true second-degree IB that p. In general, for d^{th}-degree plural IB that p, (i) every subject (in R) has the same belief that p, and (ii) every subject (in R) has the same true d^{th}-degree IB that p. Plural IB thus meets the three characteristics of shared belief.

These three characteristics of shared belief do not apply to the interactive model of mutual belief. The first characteristic applies but not the second or third; that is, although every subject believes that *p*, subjects do not believe *that they each have the same beliefs*. Subjects lack the reflexive cognizance that is required to hold the belief that they each have the same beliefs. The reason for this is to be found in the rigorous individualism of the interactive approach. Individual subjects relate to each other only as individual entities, not as members of the same group or class. But without self-awareness of themselves as members of the same group, subjects are unable to cognize that they – the members of the group – have the same beliefs. This individualism is given expression by the external universal quantification of the interactive approach: every member of set *A* is analysed by the theorist or external observer as believing that *p* (to some order of belief) but these members have no cognizance that they themselves are members of set *A*.

Mutual belief is often presented informally as 'everyone believes that everyone believes that … *p*', but this is potentially misleading about the contents of subjects' beliefs. In standard notation, mutual belief to the nth order is: $MB^n(p) \equiv B_e(p) \,\&\, B_e^2(p) \,\&\, ... \,\&\, B_e^n(p)$, where $1 \leq n < \infty$ and the subscript '*e*' is 'every subject in set *A*'. Mutual belief of the first order, $MB(p) \equiv B_e(p)$, is 'a_1 believes that *p* & a_2 believes that *p* & … & a_n believes that *p*', which is equivalent to 'every subject in *A* believes that *p*'. Mutual belief of the second order is: $MB^2(p) \equiv B_e(p) \,\&\, B_e^2(p)$. The second term, $B_e^2(p)$, is often rendered as 'everyone believes that everyone believes that *p*' but this is misleading if it is taken to imply that the distributive class, or set *A*, itself appears in the content of subjects' beliefs. The second term, $B_e^2(p)$, is equivalent to: every subject in *A* believes 'a_1 believes that *p* & a_2 believes that *p* & … & a_n believes that *p*'; it is *not* equivalent to: every subject in *A* believes 'every subject in *A* believes that *p*'. There is no internal universal quantification in the interactive hierarchy of beliefs which is given entirely in terms of external universal quantification. As an illustration of the significance of this distinction, first consider first-order mutual belief that 'every ball in box *B* is red'. If every member of *A* believes 'every ball in box *B* is red', this is equivalent to: every member of *A* believes 'ball$_1$ is red & ball$_2$ is red & … & ball$_n$ is red'. Now consider the second term in second-order mutual belief that *p*: every subject in *A* believes 'a_1 believes that *p* & a_2 believes that *p* & … & a_n believes that *p*'. Second-order mutual belief models subjects' beliefs about their beliefs in the same way that first-order mutual belief models subjects' beliefs about balls in box *B* being red. The interactive model thus *objectifies* subjects' beliefs about their beliefs by treating them (that is, subjects' beliefs about their beliefs) as structurally identical to

subjects' beliefs about physical entities; that is, subject i's belief that 'a_1 believes that p & a_2 believes that p & … & a_n believes that p' is isomorphic to subject i's belief that 'ball$_1$ is red & ball$_2$ is red & … & ball$_n$ is red'. Subjects do not have cognizance of set A so they cannot cognize that they themselves are members of set A. For this reason they cannot cognize that they each have the same beliefs. This is the fundamental explanation of why the interactive approach does not satisfy characteristics (ii) and (iii) above of shared beliefs.

This is illustrated by considering mutual belief for two subjects, i and j. First-order mutual belief evinces only characteristic (i) of shared belief because i and j believe only that p. In the case of second-order mutual belief, i and j both believe: p & $B_i(p)$ & $B_j(p)$. This is the interactive approximation of characteristic (ii), if there are no problems with indexicality in the expression for p, because i and j both believe 'i believes that p & j believes that p'. That is, i and j both believe that i believes that p and that j believes that p. With third-order mutual belief this approximation is unavailable. With third-order mutual belief both i and j believe: p & $B_i(p)$ & $B_j(p)$ & $B_i(B_j(p))$ & $B_j(B_i(p))$; that is, both i and j believe: p & $B_i(p$ & $B_j(p))$ & $B_j(p$ & $B_i(p))$. Here both i and j believe that i's beliefs and j's beliefs are different. They believe that i's and j's beliefs have different content: i believes that she herself believes 'j believes that p', whereas she believes that j believes 'i believes that p'. They also believe that i's and j's beliefs have different linguistic meaning: i and j both believe 'I believe "you believe p" and you believe "I believe p"'. These differences are exacerbated with higher orders of mutual belief. Even if there is no indexicality in the expression for p, neither i nor j believes that i's beliefs are the same as j's beliefs.[19]

The potential misunderstanding of 'everyone believes that everyone believes that p' as 'everyone in A believes "everyone in A believes that p"' corresponds to an intuitive understanding of shared belief even though it is not compatible with the interactive model. This intuitive understanding is captured by the intersubjective approach which provides a means of theorising the intuitive sense of 'everyone' in the content of a shared belief as those included in 'us'. As argued above, internal universal quantification is a manifestation of subjects' self-

---

[19] If 'every subject' in the content of a belief is taken as 'every other subject', mutual belief is construed along the lines of: 'every subject believes that p, every subject believes that every other subject believes that p, every subject believes that every other subject believes that every other subject believes that … p'. This introduces problems at the second level where i believes $(p$ & $B_j(p))$ and j believes $(p$ & $B_i(p))$. Now (ii) is not even approximated because neither subject has a belief about what i and j believe. Again, these problems are not resolved with higher orders of mutual belief.

understanding of themselves as members of a (distributive) class, and this is theorised in the intersubjective approach in terms of subjects' inclusion of themselves in the plurality, 'us'. In specific contexts subjects have cognizance of the reference group, R, and its criteria of membership, and see themselves and other subjects as members of this group. What is distinctive about shared belief is precisely that doxastic subjects' cognize each other as the other members of a reference group within which 'each of us has the same belief that p'. This also provides an interpretation of ordinary English expressions such as 'we believe that p' or 'we all believe that p', as 'each of us (in R) has the same belief that p', thus steering clear of introducing holism or collectivism into an analysis of individual subjects' beliefs.[20] This solution is not available to the interactive approach which construes 'we believe that p' or 'we all believe that p' in terms of strings of individual beliefs such as 'a_1 believes that p & a_2 believes that p & … & a_n believes that p'. It is these contrasts between the intersubjective and interactive approaches that explain why characteristics (ii) and (iii) of shared belief outlined above do not apply to mutual belief but do apply to plural intersubjective belief which, unlike mutual belief, provides a means of theorising the phenomenology of individual subjects' self-understanding that they have shared beliefs.

The intersubjective model of shared belief is *non-reductive* because plural IB that p is not explained solely in terms of individual beliefs (of whatever degree or order) that p. Individual beliefs that p are necessary but not sufficient for plural IB that p, which also requires that subjects believe 'each of us (in R) has the same belief that p' (to some degree of intersubjectivity) which in turn implies that subjects in R have cognizance of the reference group, R, and their own and others' inclusion within it. This self-understanding of subjects is in contrast with the interactive conception of mutual belief which is constructed solely out of individual beliefs that p (of whatever order) which are held by subjects who situate themselves and their beliefs solely in relation to a list of subjects and their beliefs.

Although plural IB that p is not reducible to individual subjects' beliefs that p, it is summative with respect to individual subjects' IBs that p because there is plural IB that p iff every subject in R holds the same IB that p. This distinctive combination of non-reducibility (with respect to 'each of us (in R) has the same belief that p') and summativeness (every subject holds the same IB that p) is the product of the intersubjective model's combination of internal and external

---

[20] Collective belief is not considered in this paper.

universal quantification. The following sections explain why this is an advantage in modelling shared belief.

This section has argued that the intersubjective model of plural IB theorises shared belief in terms of the standpoint of individual subjects who cognize themselves and others in terms of 'each of us'. Holding beliefs that one takes to be shared amongst the inclusive plurality, 'us', implies a notion of doxastic community within which one includes oneself. To believe that a belief is shared amongst this community involves not only the issue of whether p is true but also the issue of how subjects situate themselves or identify themselves as members of the community. This also highlights the significance of ascriptions of 'us' in defining who belongs to the same community as 'me' or is instead 'other' to 'us'. This might help to explain the 'herd' nature of many human beliefs, since holding a belief that is shared within some doxastic community involves situating oneself within or identifying with that particular community. To have a shared belief challenged is not only to have a challenge to one's own belief but also to experience a challenge to the community of which one holds oneself to be a member. This might help to explain the longevity of some false shared beliefs.

V Common belief

That something is common belief (also common knowledge) is held to be open and obvious to all concerned. Common belief is widely regarded as ubiquitous in everyday life. For example, if two people exchange glances it is common belief between them that they exchange glances; if neighbours stop to chat in the street it is common belief amongst them that they are chatting in the street; if there is a public ceremony it is common belief amongst those attending the public ceremony that they are attending such a ceremony. Furthermore, common belief is widely considered to be a vital ingredient for many social institutions and coordination processes (cf. Chwe 2001 on common knowledge). For example, the viability of a currency requires common belief in its acceptability, and coordination in driving practices and in social etiquette presupposes common belief in those practices. Of course, common belief can be mistaken. In previous times, for example, it was common belief that the Earth did not move and that some women were witches. Even when common belief is mistaken, however, the fact of common belief can play an important role in human life.

Common belief is modelled in the interactive approach in terms of an infinite version of mutual belief: $CB(p) \equiv B_e(p) \,\&\, B_e^2(p) \,\&\, \ldots$. Equivalently, common belief is defined recursively as

'every subject in A believes "p & there is common belief that p"', that is, $CB(p) \equiv B_e(p \,\&\, CB(p))$.[21] Although the fact of common belief in human society is held to be open and obvious to all concerned, there is something elusive about it in the interactive approach. As empirical human subjects do not hold or even understand infinite belief hierarchies or recursive systems, it is not clear what human subjects believe when there is common belief amongst them (cf. Ernst 2011 on common knowledge). As the evidence suggests that common belief is an essential ingredient in human life, this theoretical difficulty cannot be explained away by the argument that human subjects rarely have common beliefs. It might perhaps be suggested that low orders of mutual belief suffice for most empirical purposes, but against this suggestion it is not clear how low orders of mutual belief could suffice if it is common belief that is held to be essential in human life. In view of the logical complexity involved, some theorists in the interactive approach argue that common knowledge, and hence I take it, common belief, is ascribed to subjects who are modelled as occupying particular places within a formal system, but where subjects are not taken actually to have such knowledge or beliefs or to be able to answer questions based on them (cf. Fagin et al. 2003, pp. 8–9, 119; Aumann 1999). There is thus an ambivalence about common belief (and common knowledge) for human subjects. On the one hand, common belief is held to be an essential element of human life and distinct from mutual belief; on the other hand, given the logical complexity of common belief it is not clear what human subjects could believe, with some theorists arguing that common belief comprises only 'external' belief ascribed to subjects in a formal model.

Howsoever these issues are resolved for the interactive approach, the intersubjective approach to common belief provides a relatively simple answer to the question of what subjects believe when there is common belief amongst them, without infinite regress. In line with the argument of the previous section, according to the intersubjective approach the standard conception of common belief is too individualistic to capture the phenomenon of common belief amongst 'us' as it registers only 'I believe that you believe that I believe … and so forth'. Even with infinite belief hierarchies, subjects cannot cognize that they each have the same beliefs, so the interactive conception of common belief also fails to display characteristics (ii) and (iii) of shared belief outlined above.

---

[21] Lewis's common knowledge was constructed in terms of 'reason to believe', that is, 'everyone has reason to believe that p, everyone has reason to believe that everyone has reason to believe that p, … ' (Lewis 2002/1969), but his work has come to be seen as canonical in providing the origin of the standard interactive model of common knowledge and common belief, although it was apparently pre-dated by Friedell (1969/1967).

Common belief is possible only in a particular kind of situation, which I term a *transparent situation*. Simple examples of a transparent situation are given by two people exchanging glances, neighbours chatting in the street, and social or public ceremonies. More complex examples are given by discourse communities such as academic disciplines, political parties, religious groupings, nation states, or even all educated adults in the world. A transparent situation for subjects in R is such that, if every subject in R has the same belief that p (a belief that might be true or false) then every subject in R has the same true belief that 'each of us (in R) has the same belief that p'. This is equivalent to saying that if every subject in R has the same belief that p, then every subject in R has the same true IB that p', viz. $\bar{B}_R(p) \rightarrow \overline{IB}_R^T(p)$, where $\bar{B}_R(p)$ is 'every subject in R has the same belief that p', and $\overline{IB}_R^T(p)$ is 'every subject in R has the same true IB that p'. This assumes that every subject in R has the appropriate cognitive capacities to believe 'each of us (in R) has the same belief that p' if she believes that p. The transparency of the situation relates to subjects having the same belief that p, not to the truth of p.

This provides for the definition of a transparent situation:

> A situation is a **transparent situation** iff it is the case that, if every subject who includes herself in reference group, R, has the same belief that p, then every subject who includes herself in reference group, R, has the same (true) IB that p',
>
> that is, $\bar{B}_R(p) \rightarrow \overline{IB}_R^T(p)$.

For example, if event E is a transparent situation then, if every attendee at E has the same belief that the entertainer is walking onto the stage, then every attendee has the same belief that 'each of us (attendees at event E) has the same belief that the entertainer is walking onto the stage'. Similarly, if the entertainer is a conjuror, if every attendee has the same belief that the conjuror has just cut her assistant in half yet the assistant is alive and well, every attendee has the same belief that 'each of us (attendees at the conjuring performance) has the same belief that the conjuror has just cut her assistant in half yet the assistant is alive and well'. We might think such attendees are over-credulous but the structure of the transparent situation is the same whether p is true or false.

In a transparent situation, if subject i truly believes 'each of us (in R) has the same belief that p' *and* truly believes 'if each of us (in R) has the same belief that p, then each of us (in R) has the same true IB that p', then i validly concludes 'each of us (in R) has the same true IB that

p'. I term this inference, the *intersubjective inference*. The intersubjective inference is an inference from a true IB that *p* to a true second-degree IB that *p*. This is independent of the truth or otherwise of *p*.

That is, in a transparent situation, subject *i* who includes herself in *R* makes a valid **intersubjective inference** from a true first-degree IB that *p* to a true second-degree IB that *p* iff:

> (i) *i* holds a true first-degree IB that *p*, viz. $IB_{i/R}^{T}(p) \equiv B_{i/R}^{T}(\bar{B}p)$,
>
> (ii) *i* holds the true belief: 'if each of us (in *R*) has the same belief that *p*, then each of us (in *R*) has the same true IB that *p*', viz. $B_{i/R}^{T}(\bar{B}p \rightarrow \overline{IB}^{T}p)$,
>
> (iii) *i* concludes: 'each of us (in *R*) has the same true IB that *p*'; that is, *i* holds a true second-degree IB that *p*, viz. $IB_{i/R}^{2,T}(p) \equiv B_{i/R}^{T}(\overline{IB}^{T}p)$.[22]

To make the intersubjective inference, a subject in a transparent situation assumes: 'if each of us (in *R*) has the same belief that *p*, then each of us (in *R*) has the cognitive capacities to have the same true IB that *p*'.

The intersubjective inference provides the resources for the *intersubjective conception of common belief*, according to which there is common belief iff every subject makes the same intersubjective inference. That is, there is intersubjective common belief iff there is first-degree plural IB and every subject makes the same valid intersubjective inference, resulting in second-degree plural IB:

> There is **intersubjective common belief** that *p* amongst subjects who include themselves in reference group, *R*, iff there is first degree plural IB that *p* and every subject makes the same valid intersubjective inference, resulting in second-degree plural IB that *p*.

That is, for subjects in a transparent situation there is intersubjective common belief that *p* amongst subjects who include themselves in reference group, *R*, iff:

---

[22] The conclusion of the intersubjective inference assumes $B_i(p \rightarrow q) \rightarrow [B_i(p) \rightarrow B_i(q)]$, that is, if *i* believes $(p \rightarrow q)$ then, if *i* believes that *p*, *i* believes that *q*. Cf. Axiom K in doxastic logic. Applied to true intersubjective belief, this yields: $\bar{B}_R^T(\bar{B}p \rightarrow \overline{IB}p) \rightarrow [\bar{B}_R^T(\bar{B}p) \rightarrow \bar{B}_R^T(\overline{IB}p)]$; that is, if *i* truly believes $(\bar{B}p \rightarrow \overline{IB}p)$ then, if *i* has true first-degree IB that *p*, *i* has true second-degree IB that *p*.

(i) There is first-degree plural IB that p amongst subjects in R,

viz. $PIB_R(p) \equiv \bar{B}_R^T(\bar{B}p)$.

(ii) Every subject in R has the same true belief: 'if each of us (in R) has the same belief that p, then each of us (in R) has the same true IB that p',

viz. $\bar{B}_R^T(\bar{B}p \rightarrow \overline{IB}^T p)$.

(iii) Every subject in R concludes: 'each of us (in R) has the same true IB that p'; that is, there is second-degree plural IB that p amongst subjects in R,

viz. $PIB_R^2(p) \equiv \bar{B}_R^T(\overline{IB}^T p)$.

It was argued above that first-degree plural IB does not logically require second-degree plural IB. If, however, first-degree plural IB results in second-degree plural IB via the intersubjective inference, this is intersubjective common belief.

Intersubjective common belief (ICB) is given in (5) where $ICB_R(p)$ is 'intersubjective common belief that p amongst subjects in R':

$$ICB_R(p) \equiv [PIB_R(p) \ \& \ \bar{B}_R^T(\bar{B}p \rightarrow \overline{IB}^T p)] \rightarrow PIB_R^2(p) \tag{5}$$

It is the intersubjective inference made by every subject (in a transparent situation) that explains the phenomenological experience of openness and obviousness that is characteristic of common belief.

Returning again to the example of the team leader who has to convey information to the other members of the research team, this time the team leader informs them by announcing that p to a meeting at which all the members of the research team are present. The meeting is a transparent situation so there is intersubjective common belief that p amongst all the members of the team (assuming that the other team members believe the team leader). As commentators have observed, the significance of common belief is that information (or disinformation) is out in the open amongst a group, rather than being subject to private beliefs.

As with plural IB, intersubjective common belief can be true or false depending on whether p is true or false. If there is common belief at the conjuring event that the conjuror has just cut her assistant in half yet the assistant is alive and well, the common belief is false because it is not true that the conjuror has just cut her assistant in half, although the assistant is indeed alive and well. At the meeting of members of the research team, if the team leader's information is correct, there is true common belief amongst the members of the team.

Intersubjective common belief does not rely on an infinite hierarchy (or recursive system) of beliefs. This explains why the intersubjective model of common belief is more economical than the interactive model of common belief. As in the comparison of plural IB with mutual belief, the three characteristics of shared belief apply to intersubjective common belief although not to interactive common belief. As with plural IB, intersubjective common belief is not reducible to individual beliefs that p, although it is summative in requiring that every subject makes the same intersubjective inference. The intersubjective model of common belief thus captures subjects' self-understanding that having shared beliefs with others presupposes their self-inclusion in 'us'.

The contrast between intersubjective common belief, in (5), and the interactive approach to common belief is clear cut. The only way for the interactive approach to differentiate between mutual belief and common belief is to construe common belief in terms of an infinite hierarchy (or recursive system) of beliefs. By contrast, as the intersubjective approach differentiates between degrees of intersubjective belief and the intersubjective inference, it does not need to rely on an infinite degree of plural IB or an infinite series of intersubjective inferences, in order to differentiate between plural IB and common belief. This explains why the intersubjective approach provides an economical model of common belief that makes only modest assumptions about subjects' cognitive capacities. Of course, a transparent situation could be redefined such that subjects make an infinite series of valid intersubjective inferences, if they were to be assumed to have the cognitive capacity to engage in such infinite reasoning. But defining intersubjective common belief in terms of an infinite series of intersubjective inferences would be analytically redundant because a single intersubjective inference is sufficient to differentiate between plural IB and common belief, and to capture subjects' phenomenological experience of common belief.[23]

It is now possible to give an answer to the question of what subjects believe when there is common belief amongst them: every subject moves from first-degree IB to second-degree IB via the intersubjective inference; that is, they move from the belief that 'each of us has the same belief that p' to the belief that 'each of us has the same belief that each of us has the same belief

---

[23] Plural intersubjective knowledge (PIK) and intersubjective common knowledge (ICK) are defined analogously with their doxastic counterparts (cf. n. 10), viz. $PIK_R(p) \equiv \overline{K}_R(\overline{K}p)$; $ICK_R(p) \equiv [PIK_R(p) \ \& \ \overline{K}_R(\overline{K}p \rightarrow \overline{IK}p)] \rightarrow PIK_R^2(p)$. Hybrid forms of intersubjective belief / knowledge are also possible, for example, a subject might believe 'each of us (in R) has the same knowledge that p' or she might know 'each of us (in R) has the same belief that p'.

that p'. Any model of belief might involve an element of abstraction or logical complexity that goes beyond empirical subjects' beliefs or their understanding of their beliefs, so it is not claimed in this paper that empirical subjects do in fact in situations of common belief hold beliefs exactly as given by the intersubjective model and are aware of doing so.[24] In contrast with the interactive approach, however, the intersubjective model of common belief translates into beliefs that empirical human subjects could plausibly be interpreted as holding. The intersubjective model captures the crucial element of common belief, that subjects believe 'as each of us has the same belief that p, then each of us has *this* same belief too'. It is this single inference that explains the experience of the openness and obviousness of common belief amongst 'us'. As with mutual belief, what is missing from the interactive approach to common belief is the reflexive standpoint of subjects who see themselves as one of 'us'. Without the inclusive plurality, 'us', amongst whom subjects recognise that there is shared belief, there can be only 'mirrored' beliefs, not shared beliefs, as subjects mirror the beliefs of other subjects who in turn mirror theirs.[25]

VI Intersubjective belief: An application – Hi-Lo

As human subjects we take it for granted that we live in a shared world. The intersubjective approach proposed in this paper provides an analysis of the phenomenology of self-inclusion in terms of the standpoint of *each of us*. Earlier sections have focussed on the analytic advantages of the intersubjective approach in providing a model of shared belief, including common belief, which is both plausible and economical, and which also provides an analysis of ordinary expressions in English such as 'we believe' and 'we all believe'. This concluding section provides an illustration of how the intersubjective model provides new resources for analysing issues that have so far proved intractable or challenging using the interactive approach.

As noted above, common belief is widely understood to be important for coordination in everyday situations. But in game theory some coordination games have resisted solution even with the assumption of (interactive) common belief (common knowledge) about the game and about players' rationality (in aiming to maximize individual payoff). One of the simplest of

---

[24] What empirical human subjects actually believe when there is common belief amongst them is an empirical question that cannot be settled simply by proposing a new model of common belief.

[25] On the mirror analogy, see Lewis (2002/1969) p. 32.

these irresolvable coordination games is Hi-Lo. A two-player Hi-Lo game is illustrated in the payoff matrix in Figure 1.

Figure 1   Hi-Lo

Player 2

High   *Low*

	High	*Low*
High	*h, h*	0, 0
Low	0, 0	*l, l*

Player 1

$h > l > 0$

Player 1 and Player 2 each choose one of two actions, *High* and *Low*, and in so choosing they aim to maximize their own individual payoff. If Player 1 and Player 2 both choose *High* the outcome (*h*, *h*) is better for both than if they both choose *Low* (*l*, *l*) because $h > l$. The worst outcome for both players is if one chooses *High* and the other chooses *Low* as payoffs are zero. Both (*High*, *High*) and (*Low*, *Low*) are Nash equilibria (action profiles in which each player's action is best individually given the other's action). The question is: should a player choose *High* or *Low*?

Intuitively it seems obvious that the answer is '*High*'. For standard game theory there is no determinate answer. The reasoning goes like this: Player 1 should choose *High* if Player 2 chooses *High*, but Player 2 should choose *High* if Player 1 chooses *High*, but Player 1 should choose *High* if Player 2 chooses *High*, …, and so on, in endless regress. In other words, it is common belief for the players that they reason in terms of 'I' and 'you' as follows: 'I should choose *High* if you choose *High*, but you should choose *High* if I choose *High*, but I should choose *Hig*h if you choose *High*, … ; on the other hand, I should choose *Low* if you choose *Low*, but you should choose *Low* if I choose *Low*, … '. There is thus no game-theoretic reason for choosing *High* over *Low* or *Low* over *High*. There is no determinate answer as to which action a player should choose.

In spite of the simplicity of Hi-Lo it raises some fundamental issues for standard game theory which can hardly be addressed here. Suffice it to say that the intersubjective approach rejects analysis in terms of common belief denominated in terms of 'I' and 'you'. Hi-Lo is a shared

game situation for the players which they cognize in terms of 'each of us'; for example, 'each of us (Players 1 & 2) is playing Hi-Lo'. It is also a transparent situation, so there is intersubjective common belief amongst the players about the game and that 'each of us aims to maximize individual payoff'. It is intuitively obvious to players that 'if each of us chooses *High*, the individual payoff for each of us is maximized'; and this is common belief. Accordingly, it is common belief that 'each of us should choose *High*'. By the Principle of Intersubjective Inclusion, if 'each of us should choose *High*' then for both players it follows that 'I should choose *High*'. This is common belief too. Both players choose *High* in common belief that each of them should choose *High*.

The intersubjective approach theorizes subjects' cognizance of their shared world in terms of their self-inclusion in 'each of us', not in terms of 'I' and 'you'. This cognizance informs subjects' beliefs about their shared beliefs which are also expressed in terms of 'each of us'. As Hi-Lo illustrates, the intersubjective approach has the potential to produce new results in shared contexts where reductive approaches are too individualistic to capture the crucial element of sharedness. This potential is not restricted to situations where there is a congruence of individual interests (as in Hi-Lo) but is applicable to a diverse range of shared contexts.

References

Aumann, Robert J. (1976) 'Agreeing to disagree', *The Annals of Statistics*, 4: 1236–1239.

—— (1999) 'Interactive epistemology I: Knowledge', *International Journal of Game Theory*, 28: 263–300.

Braun, David, (2016) 'Indexicals', *The Stanford Encyclopedia of Philosophy* (Fall 2016 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2016/entries/indexicals/>.

Chwe, Michael Suk-Young (2001) *Rational Ritual: Culture, Coordination, and Common Knowledge*, Princeton University Press.

Ernst, Zachary (2011) 'What is common knowledge?', *Episteme*, 8: 209–226.

Fagin, Ronald, Halpern, Joseph Y., Moses, Yoram, and Vardi, Moshe Y. (2003) *Reasoning About Knowledge*, MIT Press; first edition 1995.

Friedell, Morris F. (1969) 'On the structure of shared awareness', *Behavioral Science*, 14: 28–39; Working Paper 27, Centre for Research on Social Organization, University of Michigan, April 1967.

Kaplan, David (1989) 'Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives, and Other Indexicals', in *Themes from Kaplan*, Joseph Almoog, John Perry and Howard Wettstein (eds), Oxford University Press, pp. 481–563.

Lewis, David (2002) *Convention: A Philosophical Study*, Blackwell Publishers; first published, 1969, Harvard University Press.

Vanderschraaf, Peter and Sillari, Giacomo (2014) 'Common Knowledge', *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2014/entries/common-knowledge/>.

Weirich, Paul (2011) 'Introduction: Interactive epistemology', *Episteme*, 8: 201–208.