investigation of multiperson rationality and presentation effects. *Game Econ. Beh.* 6, 445–468

14 Stahl, D.O. and Wilson, P.W. (1995) On players' models of other players: theory and experimental evidence. *Game Econ. Behav.* 10, 218–254

15 Colman, A.M. and Stirk, J.A. (1998) Stackelberg reasoning in mixed-motive games: an experimental investigation. *J. Econ. Psychol.* 19, 279–293

16 Nagel, R. (1995) Unraveling in guessing games: an experimental study. *Am. Econ. Rev.* 85, 1313–1326

PII: S1364-6613(02)00006-2

Research Focus Response

Two paradigms for depth of strategic reasoning in games

Response to Colman

Jun Zhang and Trey Hedden

Department of Psychology, University of Michigan, Ann Arbor, MI 48109, USA

We contrast two approaches to probing the depth of theory-of-mind (ToM) reasoning by adults in matrix game settings: our own and that of Stahl and Wilson.

As Colman points out [1], although recursive reasoning in games is an essential element of normative game theory, empirical investigations of the sophistication of human players demonstrate that relatively shallow orders of reasoning are more usually used. Strategic reasoning in matrix games in adults using theory-of-mind (ToM) has been probed by Stahl and Wilson [2] and more recently by us [3]. Important differences in these two approaches hold implications for a descriptive theory of bounded rationality.

The Stahl–Wilson paradigm

Stahl and Wilson [2] asked participants to play a series of games against a group of opponents. No feedback was given until all games were completed, but players were allowed to reconsider choices made for previous games. In some cases, inferences could be made based on an analysis of dominance relationships between strategies. A player's depth of reasoning was determined by the extent to which dominated strategies were iteratively eliminated through perspective-taking. Research probing ToM reasoning through dominance relationships of payoffs in games was previously conducted by Perner [4].

This approach has the advantages that: (1) because no belief prompting was given, the depth of reasoning achieved by participants was spontaneous; and (2) all orders of reasoning were operationally defined, although an individual player's strategy was determinable only through statistical modeling.

However, it is not clear that the Stahl–Wilson paradigm probes the order of reasoning players are fully capable of engaging in, as opposed to beliefs about how their opponents will behave. The strategy used by a player will depend upon an estimate of the distribution of strategies used by all other players. As a consequence, some second-order players might revert to lower-order reasoning because of uncertainty about the sophistication of their opponents. It remains to be seen whether a manipulation of background information about opponents' level of intelligence, as performed in Hedden and Zhang [3], would change the distribution of players of various orders.

Although reconsideration of choices masked any learning or sequential effects, it is debatable whether players would exhibit a uniform strategy to Nash-solvable and to dominance-solvable games, because their strategic analyses differ. This is further complicated by the fact that these two kinds of games may well relate to depth of ToM in different and multiple ways. For example, *defect–defect* is the solution to the standard Prisoner's Dilemma game, in the sense of dominance (in addition to the sense of Nash) that requires only zeroth- and first-order consideration, but mutual cooperation becomes a Nash equilibrium under second-level meta-game analysis involving second- and third-order ToM reasoning between the two players [5].

The Hedden–Zhang paradigm

In contrast to Stahl and Wilson, participants in Hedden and Zhang [3] were pitted against a single opponent (a confederate programmed to use either a zeroth- or first-order strategy, or to switch strategies between two consecutive blocks). Hence, players were modeling a particular individual as opposed to estimating the distribution of strategic sophistication in a population. Further, players received feedback on each trial, allowing them to determine the opponent's strategy and to update their ToM model on-line. The availability of reaction-time measures provided corroborative evidence about participants' engagement in shallow and deep reasoning in individual games. This paradigm therefore investigates dynamic changes in the depth of reasoning resulting from progressive interaction, as well as providing an estimate of participants' initial default strategy. For most players, this default was based on first-order reasoning.

Colman [1] notes that drawbacks of this paradigm include that it: (1) relies on a particular operational definition of a zeroth-order opponent's behavior; and (2) involves belief prompting. Both are legitimate concerns.[1] However, (1) would underestimate the probability of

Corresponding author: Jun Zhang (junz@umich.edu).

[1] Colman's concern about the motivation of participants when no monetary incentive is involved falls more in the line of standard practices of economists and psychologists. Points earned by opponents were not tallied and displayed in our experiments [3], and care was taken to eliminate participants based on their performance on catch trials incorporated into the experimental design.

first-order reasoning by players, and (2) would overestimate the percentage of second-order reasoners. Although they bias the data, both argue in favor of the conclusion that human adults reason in a boundedly rational fashion with shallow application of ToM orders.

In conclusion, both paradigms have advantages and disadvantages, so care must be taken in interpreting experimental findings from each. Stahl and Wilson's paradigm is free of belief prompting, tests spontaneous models based upon estimates about a population of opponents, but might overestimate the percentage of first-order reasoners by virtue of uncertainty about and lack of interaction with opponents. Hedden and Zhang's paradigm probes initial and dynamic modeling of an opponent through belief-prompting, but might underestimate the preponderance of first-order reasoning.

References

1 Colman, A.M. (2003) Depth of strategic reasoning in games. *Trends Cogn. Sci.* 7, 1–2
2 Stahl, D.O. and Wilson, P.W. (1995) On player's models of other players: theory and experimental evidence. *Games Econ. Behav.* 10, 218–254
3 Hedden, T. and Zhang, J. (2002) What do you think I think you think?: strategic reasoning in matrix games. *Cognition* 85, 1–26
4 Perner, J. (1979) Young children's preoccupation with their own payoffs in strategic analysis of 2×2 games. *Dev. Psychol.* 15, 204–213
5 Howard, N. (1971) *Paradoxes of Rationality*, MIT Press

PII: S1364-6613(02)00007-4

Research Focus

Space is special in Sign

Ruth Campbell[1] and Bencie Woll[2]

[1]Department of Human Communication Science, University College London, Chandler House, 2 Wakefield Street, London WC1N 1PG, UK
[2]Department of Language and Communication Science, City University London, London EC1V 0HB, UK

Following groundbreaking work by linguists and cognitive scientists over the past thirty years, it is now generally accepted that sign languages of the deaf, such as ASL (American Sign Language) or BSL (British Sign Language), are structured and processed in a similar manner to spoken languages. The one striking difference is that they operate in a wholly non-auditory, visuospatial medium. How does the medium impact on language processing itself?

Behavioural studies of focal lesion patients with ASL as a first language [1,2] suggest that the impact is minimal. Although space is the medium in which sign language is expressed, spatial debilities following right-hemisphere (RH) impairment have a minor impact on *linguistic* processing in ASL. There is little indication that the left hemisphere (LH) of signers does anything especially 'spatial' compared with the LH of speakers. However, since the mid-1980s Helen Neville and her colleagues, who have pioneered brain imaging studies of ASL, have consistently reported relatively greater contributions of RH processing to sign language than might occur for processing English (e.g. [3,4]). These findings have generated a good deal of debate: Although they demonstrated ASL processing makes use of RH systems, it was not clear to what extent these were specific to sign language [5,6].

Looking for space in sign

Now an fMRi and a PET study of two groups of native signers – one of BSL and one of ASL[1] – cast some light on the question of how signed languages may make use of cortical systems specialised for spatial processing. Although all sign languages make use of space, there are some constructions in which space is used in a special way. Both studies have focussed on these constructions. MacSweeney *et al.* [7], in a fMRI study of BSL users, contrasted the comprehension of two types of BSL utterance: topographic and non-topographic. Topographic utterances make use of space in an isomorphic and analogue fashion. For topographic sentences, there is a schematic correspondence between the location and movement of the hands in signing space (handshapes represent objects or persons) and the position and movement of objects in the real (or imagined) world. For example, in the BSL sentence translated as '*The car turned left and hit the lorry*', handshape modifications (called 'classifiers' by sign linguists) represent the two vehicles, whereas the start and end points of the hand's actions for '*car*', and the pattern of movement between them, directly represent moving forward, left and stopping (for more examples, see http://www.ich.ucl.ac.uk/macsweeney/jocnstimuli).

Emmorey *et al.* [8] explored space in a different way, adapting a PET paradigm developed by Damasio *et al.* [9] that required signed descriptions of pictured spatial relations (e.g. '*cup on table*'). The most common way to 'name' such object relations is by the spatial arrangement of classifier signs. To express '*cup on table*' the hand representing the cup (a curved hand) is placed on top of the hand representing the table (a flat hand, palm down). In contrast, English speakers use lexical prepositions such as ON, IN, etc. to indicate spatial relationships. Unlike prepositions, the sentences with classifier signs signal

Corresponding author: Ruth Campbell (r.campbell@ucl.ac.uk).
[1] ASL and BSL are historically unrelated and mutually unintelligible.