

Dear Author,

Here are the proofs of your article.

- You can submit your corrections **online**, via **e-mail** or by **fax**.
- For **online** submission please insert your corrections in the online correction form. Always indicate the line number to which the correction refers.
- You can also insert your corrections in the proof PDF and **email** the annotated PDF.
- For fax submission, please ensure that your corrections are clearly legible. Use a fine black pen and write the correction in the margin, not too close to the edge of the page.
- Remember to note the **journal title**, **article number**, and **your name** when sending your response via e-mail or fax.
- **Check** the metadata sheet to make sure that the header information, especially author names and the corresponding affiliations are correctly shown.
- **Check** the questions that may have arisen during copy editing and insert your answers/ corrections.
- **Check** that the text is complete and that all figures, tables and their legends are included. Also check the accuracy of special characters, equations, and electronic supplementary material if applicable. If necessary refer to the *Edited manuscript*.
- The publication of inaccurate data such as dosages and units can have serious consequences. Please take particular care that all such details are correct.
- Please **do not** make changes that involve only matters of style. We have generally introduced forms that follow the journal's style. Substantial changes in content, e.g., new results, corrected values, title and authorship are not allowed without the approval of the responsible editor. In such a case, please contact the Editorial Office and return his/her consent together with the proof.
- If we do not receive your corrections **within 48 hours**, we will send you a reminder.
- Your article will be published **Online First** approximately one week after receipt of your corrected proofs. This is the **official first publication** citable with the DOI. **Further changes are, therefore, not possible.**
- The **printed version** will follow in a forthcoming issue.

Please note

After online publication, subscribers (personal/institutional) to this journal will have access to the complete article via the DOI using the URL: [http://dx.doi.org/\[DOI\]](http://dx.doi.org/[DOI]).

If you would like to know when your article has been published online, take advantage of our free alert service. For registration and further information go to: <http://www.springerlink.com>.

Due to the electronic nature of the procedure, the manuscript and the original figures will only be returned to you on special request. When you return your corrections, please inform us if you would like to have these documents returned.

ELECTRONIC REPRINT ORDER FORM

After publication of your journal article, electronic (PDF) reprints may be purchased by arrangement with Springer and Aries Systems Corporation.

The PDF file you will receive will be protected with a copyright system called DocuRights®. Purchasing 50 reprints will enable you to redistribute the PDF file to up to 50 computers. You may distribute your allotted number of PDFs as you wish; for example, you may send it out via e-mail or post it to your website. You will be able to print five (5) copies of your article from each one of the PDF reprints.

Please type or print carefully. Fill out each item completely.

1. Your name: _____
 Your e-mail address: _____
 Your phone number: _____
 Your fax number: _____
2. Journal title (vol, iss, pp): _____
3. Article title: _____
4. Article author(s): _____
5. How many PDF reprints do you want? _____
6. Please refer to the pricing chart below to calculate the cost of your order.

Number of PDF reprints	Cost (in U.S. dollars)
50	\$200
100	\$275
150	\$325
200	\$350

NOTE: Prices shown apply only to orders submitted by individual article authors or editors. Commercial orders must be directed to the Publisher.

All orders must be prepaid. Payments must be made in one of the following forms:

- a check drawn on a U.S. bank
- an international money order
- Visa, MasterCard, or American Express (no other credit cards can be accepted)

PAYMENT (type or print carefully):

Amount of check enclosed: _____ (payable to Aries Systems Corporation)

VISA _____

MasterCard _____

American Express _____

Expiration date: _____ Signature: _____

Print and send this form with payment information to:

Aries Systems Corporation
 200 Sutton Street
 North Andover, Massachusetts 01845
 Attn.: Electronic Reprints
 — OR —
 Fax this to Aries at: 978-975-3811

Your PDF reprint file will be sent to the above e-mail address. If you have any questions about your order, or if you need technical support, please contact: support@docurights.com

For subscriptions and to see all of our other products and services, visit the Springer website at:

<http://www.springeronline.com>

Metadata of the article that will be visualized in OnlineFirst

Please note: Images will appear in color online but will be printed in black and white.

ArticleTitle	Metamemory as evidence of animal consciousness: the type that does the trick	
Article Sub-Title		
Article CopyRight	The Author(s) (This will be the copyright line in the final PDF)	
Journal Name	Biology & Philosophy	
Corresponding Author	Family Name	Shea
	Particle	
	Given Name	Nicholas
	Suffix	
	Division	Faculty of Philosophy
	Organization	University of Oxford
	Address	10 Merton Street, OX1 4JJ, Oxford, UK
	Division	
	Organization	Somerville College
	Address	Oxford, UK
	Email	nicholas.shea@philosophy.ox.ac.uk
Author	Family Name	Heyes
	Particle	
	Given Name	Cecilia
	Suffix	
	Division	
	Organization	All Souls College
	Address	Oxford, UK
	Email	
Schedule	Received	25 September 2008
	Revised	
	Accepted	18 May 2009
Abstract	<p>The question of whether non-human animals are conscious is of fundamental importance. There are already good reasons to think that many are, based on evolutionary continuity and other considerations. However, the hypothesis is notoriously resistant to direct empirical test. Numerous studies have shown behaviour in animals analogous to consciously-produced human behaviour. Fewer probe whether the same mechanisms are in use. One promising line of evidence about consciousness in other animals derives from experiments on metamemory. A study by Hampton (Proc Natl Acad Sci USA 98(9):5359–5362, 2001) suggests that at least one rhesus macaque can use metamemory to predict whether it would itself succeed on a delayed matching-to-sample task. Since it is not plausible that mere meta-representation requires consciousness, Hampton's study invites an important question: what kind of metamemory <i>is</i> good evidence for consciousness? This paper argues that if it were found that an animal had a memory trace which allowed it to use information about a past perceptual stimulus to inform a range of different behaviours, that would indeed be good evidence that the animal was conscious. That functional characterisation can be tested by investigating whether successful performance on one metamemory task transfers to a range of new tasks. The paper goes on to argue that thinking about animal consciousness in this way helps in formulating a more precise functional characterisation of the mechanisms of conscious awareness.</p>	
Keywords (separated by '-')	Animal consciousness - Metamemory - Phenomenal consciousness - Meta-representation - Higher order theories of consciousness - Global availability - Access consciousness	

Footnote Information

Nicholas Shea and Cecilia Heyes contributed equally to this work and the order of their names is arbitrary.

Journal: 10539
Article: 9171

Author Query Form

**Please ensure you fill out your response to the queries raised below
and return this form along with your corrections**

Dear Author,

During the preparation of your manuscript for typesetting, some questions have arisen. These are listed below. Please check your typeset proof carefully and mark any corrections in the margin of the proof or compile them as a separate list. This form should then be returned with your marked proof/list of corrections to spr_corrections1@sps.co.in

Disk use

In some instances we may be unable to process the electronic file of your article and/or artwork. In that case we have, for efficiency reasons, proceeded by using the hard copy of your manuscript. If this is the case the reasons are indicated below:

- Disk damaged Incompatible file format LaTeX file for non-LaTeX journal
 Virus infected Discrepancies between electronic file and (peer-reviewed, therefore definitive) hard copy
 Other:

We have proceeded as follows:

- Manuscript scanned Manuscript keyed in Artwork scanned
 Files only partly used (parts processed differently:

Bibliography

If discrepancies were noted between the literature list and the text references, the following may apply:

- The references listed below were noted in the text but appear to be missing from your literature list. Please complete the list or remove the references from the text.
 Uncited references: This section comprises references that occur in the reference list but not in the body of the text. Please position each reference in the text or delete it. Any reference not dealt with will be retained in this section.

Queries and/or remarks

Section/paragraph	Details required	Author's response
Front matter	Please check and confirm the author names and their respective affiliations are correctly identified.	
Body matter	Please confirm the section headings are correctly identified.	
Back matter	Reference Hampton et al. (2004) is given in the list but not in the text. Please provide the citation in th text.	
	Please update the reference Carruthers (2007) if possible.	

1

2 **Metamemory as evidence of animal consciousness:**
3 **the type that does the trick**

4 **Nicholas Shea · Cecilia Heyes**

5 Received: 25 September 2008 / Accepted: 18 May 2009
6 © The Author(s) 2009. This article is published with open access at Springerlink.com

7 **Abstract** The question of whether non-human animals are conscious is of
8 fundamental importance. There are already good reasons to think that many are,
9 based on evolutionary continuity and other considerations. However, the hypothesis
10 is notoriously resistant to direct empirical test. Numerous studies have shown
11 behaviour in animals analogous to consciously-produced human behaviour. Fewer
12 probe whether the same mechanisms are in use. One promising line of evidence
13 about consciousness in other animals derives from experiments on metamemory. A
14 study by Hampton (Proc Natl Acad Sci USA 98(9):5359–5362, 2001) suggests that
15 at least one rhesus macaque can use metamemory to predict whether it would itself
16 succeed on a delayed matching-to-sample task. Since it is not plausible that mere
17 meta-representation requires consciousness, Hampton's study invites an important
18 question: what kind of metamemory *is* good evidence for consciousness? This paper
19 argues that if it were found that an animal had a memory trace which allowed it to
20 use information about a past perceptual stimulus to inform a range of different
21 behaviours, that would indeed be good evidence that the animal was conscious. That
22 functional characterisation can be tested by investigating whether successful per-
23 formance on one metamemory task transfers to a range of new tasks. The paper goes
24 on to argue that thinking about animal consciousness in this way helps in

A1 Nicholas Shea and Cecilia Heyes contributed equally to this work and the order of their names is
arbitrary.

A2 N. Shea (✉)
A3 Faculty of Philosophy, University of Oxford, 10 Merton Street, Oxford OX1 4JJ, UK
A4 e-mail: nicholas.shea@philosophy.ox.ac.uk

A5 N. Shea
A6 Somerville College, Oxford, UK

A7 C. Heyes
A8 All Souls College, Oxford, UK

25 formulating a more precise functional characterisation of the mechanisms of con-
 26 scious awareness.

27
 28 **Keywords** Animal consciousness · Metamemory · Phenomenal consciousness ·
 29 Meta-representation · Higher order theories of consciousness · Global availability ·
 30 Access consciousness

32 **Investigating animal consciousness: why? How?**

33 There is a wealth of research on animals' metacognitive abilities. Some experiments
 34 are interpreted as furnishing direct empirical evidence of consciousness in animal
 35 subjects. Although evolutionary and neurological continuity give us good reason to
 36 think that some other animals are conscious, it is notoriously difficult to test that
 37 hypothesis directly, or to tell how far consciousness extends into the animal
 38 kingdom. We aim to show that conclusions about animal consciousness can be drawn
 39 from experiments on metacognition. Our focus is metamemory: an individual's
 40 ability to keep track of whether she accurately remembers a stimulus. We take
 41 metamemory as an illustrative case. It is not the only way in which conclusions about
 42 animal consciousness can be based on experimental observations, but by working
 43 through various methodological and philosophical objections in detail in this one
 44 case, we hope to demonstrate the merits of the broader methodology for investigating
 45 consciousness that we propose.

46 Since our focus is on evidence for consciousness, we do not aim to review the
 47 large comparative literature on metamemory, let alone metacognition in general. Of
 48 the many sorts of metamemory that have been studied, we are interested in the type
 49 of metamemory that can play an additional role, forming a plausible basis for
 50 inferences about consciousness. Which type of metamemory is indeed good
 51 evidence for consciousness—which type will do the trick?

52 To do the trick, the ability must be characterised in non-consciousness-involving
 53 terms *C*, in a way that makes it plausible that a subject's meeting condition *C* in
 54 relation to a perceptual stimulus is good evidence that they consciously remember it.
 55 Testing for condition *C* will then be one empirically-tractable way to probe whether
 56 other animals are conscious. Even those who reject higher order thought as
 57 necessary for consciousness should accept that some type of meta-representation
 58 can be evidence of consciousness. This paper addresses the question: what variety of
 59 meta-representation is suited to playing that evidential role?

60 The target of our investigation is phenomenal consciousness—the “what it's
 61 like”-ness of our mental lives. When we reflect on consciousness from the first
 62 person perspective, it can seem as if explaining and investigating it further is
 63 intractable. The logic of the approach taken here is to focus on what conscious
 64 experience does for us—to look for ways of characterising its functional profile. For
 65 example, there is evidence that there are two different ways of forming an
 66 association between a tone and a puff of air to the eye so that the tone comes to
 67 cause an eye blink: ‘delay conditioning’ and ‘trace conditioning’. It seems that trace
 68 conditioning requires consciousness, whereas delay conditioning does not. In delay

69 conditioning a puff of air to the eye is administered during the occurrence of a tone
 70 (after the start of the tone, hence ‘delay’ conditioning). Delay conditioning
 71 dissociates from awareness of the contingency between tone and air puff (Perruchet
 72 1985). By contrast, it seems that trace conditioning—where the air puff occurs
 73 shortly after the tone has stopped—correlates with subjects’ conscious awareness of
 74 the contingency (Clark et al. 2001; Clark and Squire 1998; Perruchet et al. 2006). If
 75 it were established that only trace conditioning requires consciousness, then the
 76 presence or absence of trace conditioning, and of the mechanisms which underlie it,
 77 could be used as evidence as to whether other animals are conscious. When
 78 following this method, it is important that the mechanism of trace conditioning in
 79 humans be characterised in detail: its functional profile, the brain mechanisms
 80 involved, modes of intervening on or interfering with those mechanisms, etc. Such a
 81 detailed characterisation C_1 of the mechanism of trace conditioning is much richer
 82 than the bare observation that trace conditioning appears to correlate with verbal
 83 report of the contingency. The process of testing whether C_1 is present in other
 84 animals is correlatively more empirically tractable (and falsifiable).

85 We call a task ‘consciousness-involving’ if humans’ performance of the task, or
 86 their performance of the task in a particular way, correlates with their being
 87 conscious of the task-relevant parameters, as indexed by subjects’ introspective and
 88 environmental reports. We can study a range of consciousness-involving tasks (Jack
 89 and Shallice 2001). Given thorough investigation, the mechanism deployed in each
 90 consciousness-involving task can be characterised in detail: C_1, C_2, \dots, C_n . Each
 91 such characterisation is then susceptible to independent investigation in animals,
 92 without relying on verbal report, to see which other animals have the C_i mechanism.
 93 The purpose of this paper is to arrive at such detailed characterisation in the case of
 94 metamemory.

95 We should distinguish three types of potentially conscious state. First, there is the
 96 online visual perception of a stimulus. Second, there is visual recall of a recent past
 97 stimulus—the kind of state you are in when you shut your eyes and visualise the
 98 scene you have just been looking at. Third, there is metamemory: some kind of
 99 representation of your own visual memory. We ask whether states of the second
 100 kind are conscious, and focus on whether states of the third kind are good evidence
 101 of such consciousness. Information about an immediately past stimulus may be held
 102 online without being conscious. The claim we are considering is that there is a type
 103 of metamemory that correlates with the perceptual memory trace being conscious.
 104 Information about an immediately past perceptual stimulus is clearly a form of
 105 memory in the broadest sense. Where it falls in relation to the standard taxonomy of
 106 memory partly depends upon whether it is conscious. If so, it would be explicit
 107 rather than implicit. It would also be declarative rather than procedural, although
 108 that is a distinction that is usually applied to long term memory, rather than the short
 109 term memory involved in keeping information about an immediately past perceptual
 110 trace online. It is also episodic in character. Indeed, debates about whether long term
 111 episodic memory is evidence for consciousness in other species (e.g., Tulving 2005)
 112 have a similar structure to the issues considered here: although experiments may
 113 uncover behaviour that depends upon information about the time and place of some
 114 particular event in the animal’s past experience, the further question arises about the

115 circumstances in which the use of such episodic information provides evidence of
 116 consciousness. The process we engage in here of examining the metamemory
 117 literature with an eye to evidence about consciousness needs to be repeated for other
 118 tasks, like those involving episodic memory, that might also allow for direct
 119 empirical testing of consciousness in animals.

120 We will not discuss the neural mechanisms of human consciousness-involving
 121 metamemory, although they may be an important part of the story, but will instead
 122 aim at a broadly functional characterisation that can be carried across from humans
 123 to other animals. We call this condition *C*. Our aim is to formulate the condition *C*
 124 appropriate to consciousness-involving metamemory: what species of metamemory
 125 goes with a human subject's having a conscious perceptual memory of a stimulus?
 126 Finding that other animals do indeed satisfy such a detailed functional condition, in
 127 some circumstances, would then be good evidence that they were conscious in those
 128 circumstances; evidence that could be further reinforced by data about neural
 129 mechanisms (which there is not space to discuss here).

130 In formulating our condition *C*, we must walk a narrow ridge between tempting
 131 mistakes of opposite kinds. On the one hand we might formulate a condition which
 132 is in fact met by unconscious systems. That danger can be addressed by rigorous
 133 studies in people to ensure that the presence or absence of our proposed condition *C*
 134 does in fact correlate with the presence or absence of consciousness as measured by
 135 verbal report. However, while we concentrate on avoiding the abyss of the
 136 unconscious on one side, we may stray onto the comfortable slopes on the other
 137 where consciousness is a decidedly human-only phenomenon. That is, to be sure
 138 people only meet condition *C* when they are conscious we may formulate a
 139 condition that is too strong, which we know only humans can meet, effectively
 140 presupposing that animals cannot be conscious. For example, we might build verbal
 141 report or its equivalent into condition *C*. The discipline of formulating a condition *C*
 142 for which animals can be tested empirically should help to avoid settling on a
 143 condition whose connection with consciousness is exemplified only in humans. That
 144 is our objective. We aim to formulate a condition *C*, associated with metamemory
 145 tasks, for which non-human animals can be tested. The animal focus serves to ward
 146 us off the comfortable slopes of anthropocentrism.

147 It follows from the logic of our approach that finding behaviour in animals which
 148 is analogous to consciously-produced human behaviour has little forensic merit.
 149 Showing that an animal can solve a problem that a human would solve using
 150 metamemory casts little light on whether the animal is conscious. Experiments must
 151 test whether humans and animals deploy the same mechanisms, our focus here
 152 being on a functional characterisation of those mechanisms. In section "[Animal](#)
 153 [data](#)" we give examples of work on metamemory that has moved towards this more
 154 stringent objective. In section "[Meta-memory: high level meta-representation](#)" we
 155 specify the type of metamemory which would be good evidence for consciousness
 156 and set out how it can be tested for in animals.

157 **Animal data**

158 Cowey and Stoerig (1995), Stoerig et al. (2002)

159 From the extensive literature on metacognition in non-human animals (Smith et al.
160 2003), we select two experimental paradigms, each to illustrate a particular point.
161 The first, which we discuss in this section, was deployed by Cowey and Stoerig in
162 an elegant series of experiments on blindsight in monkeys (Cowey and Stoerig
163 1995, 1997; Stoerig et al. 2002). The second, discussed in the next section, is a
164 memory discrimination procedure used by Hampton (2001) to provide evidence of
165 meta-representation in a rhesus monkey. Both paradigms are based on the
166 'commentary key' method devised by Weiskrantz (1986, 1995).

167 Cowey and Stoerig studied monkeys with unilateral lesions of the primary visual
168 cortex comparable to those which, in humans, give rise to blindsight—voluntary
169 responding to visual stimuli in the absence of phenomenal consciousness. These
170 lesioned animals were compared with intact controls on two successive tasks. In the
171 first, 'localisation' task, the monkeys were rewarded with food for touching the
172 visual target location, and the test stimuli were presented equally often in the right
173 hemifield, where one would expect lesion-induced impairment, and in the left
174 hemifield, where one would expect performance to be unaffected by the lesion. The
175 results indicated that, at appropriate stimulus intensities, the lesioned animals could
176 localise the stimuli presented to their right 'blind' field with almost 100% accuracy.
177 The second, 'detection' task introduced the commentary key. In 50% of trials during
178 initial training on this detection task a visual target was presented in the normal field
179 and the monkey was rewarded if it touched the target location. The other 50% of
180 trials were blanks, i.e., no target was presented, and the animal was rewarded if
181 it touched a box stimulus that was constantly present on the computer screen.
182 According to the logic of the commentary key method, touching this box constituted
183 a report by the animal that it had not seen a visual stimulus in that trial. Once this
184 discrimination had been mastered—once the animals were reliably touching the
185 target on target trials and the box on blank trials—visual targets in the right 'blind'
186 field began to be presented in 5% of trials. In these crucial probe trials reward was
187 programmed for delivery whether the animal touched the probe or the box. The
188 result was that the normal monkey consistently touched the probe, but the lesioned
189 animals nearly always (92–98% of trials) touched the box. So, in combination, the
190 two tasks showed that, when reward depends on it, monkeys with striate cortex
191 lesions can localise visual stimuli in the 'blind' hemifield, but that when they have
192 the option of getting reward without localisation, they act as they have learned to do
193 when no stimulus was presented.

194 Cowey and Stoerig's findings show that, if monkeys are conscious, they exhibit
195 blindsight in much the same way as human subjects. But that is to make the
196 (plausible) assumption that some non-human animals can be conscious, not to test it.
197 Cowey and Stoerig's studies do not demonstrate, or seek to demonstrate, that intact
198 monkeys are perceptually conscious of the visual stimuli to which they respond. This
199 is a perfectly reasonable assumption in the sense that it accords with most people's
200 intuitions, and it is put to good scientific use in their research. It is used to test a



201 'continuity' hypothesis, the idea that striate cortex lesions have the same effects in
 202 humans and monkeys, against an alternative 'encephalization' hypothesis, which
 203 suggests that hominid evolution involved migration of visual function within the
 204 brain, and therefore that these lesions will have different effects in the two species.
 205 *If* one assumes that intact monkeys are conscious of the visual stimuli to which
 206 they respond, then their results support the blindsight hypothesis, with loss of
 207 consciousness following striate cortex lesions. However, if one questions this
 208 assumption, in the way that is necessary when the purpose of enquiry is to find out
 209 whether animals are conscious, its justification turns out to depend, not on careful
 210 functional analysis of visual perception, but on reasoning by analogy from one's
 211 own case (Heyes 2008). When I respond to visual stimuli I tend to be conscious of
 212 them, therefore when a monkey responds to similar stimuli under comparable
 213 conditions, I assume that he is also conscious of them. Below we argue that a more
 214 secure inference should be based on obtaining a detailed functional characterisation
 215 of how humans perform the task when they do so in a consciousness-involving way,
 216 and carrying that over as the basis of tests in other animals.

217 Our principal concern in this section has been to point out that Cowey and
 218 Stoerig's work, and other research that makes good scientific use of the assumption
 219 that nonhuman animals are conscious, does not furnish strong evidence that other
 220 animals are conscious.

221 Hampton (2001)

222 Hampton (2001, Experiment 3) used a memory discrimination task to produce good
 223 evidence for meta-representation in rhesus monkeys. Although he disclaimed any
 224 attempt to be studying the subjective experiences of his animal subjects (p. 5359),
 225 claiming that the *experiences* associated with remembering cannot be studied in
 226 non-human animals (p. 5362), we argue that Hampton's method can form the basis
 227 of experiments that would furnish evidence about this deeper issue.

228 At the beginning of each trial in Hampton's procedure, the monkey was shown
 229 one of four pictures on a computer screen (a new set each day). After picture
 230 presentation, there was a delay of variable duration (12.5–200 s), in which the
 231 screen was blank. After the delay, the monkey was usually required to touch one of
 232 two flags on the screen. Touching the 'test flag' resulted in the monkey being
 233 presented with a display containing all four pictures. If he selected from this array
 234 the picture he had seen at the beginning of the trial, he received a preferred reward, a
 235 peanut. By touching the other 'escape flag' the monkey could avoid the test but be
 236 sure of a lesser reward, a pellet of ordinary primate food (Fig. 1).

237 The result of the experiment was that the frequency with which the monkey
 238 chose the escape flag over the test flag increased with the duration of the delay after
 239 the original picture was presented. This pattern is consistent with the use of meta-
 240 representation by the monkey of its perceptual memory. That is, because memories
 241 fade over time, one would expect the probability of choosing the escape key to
 242 increase with delay if the monkey's decision whether to press the test key or the
 243 escape key depended on the strength of an internal representation of the sample
 244 stimulus. However, this relationship between choice of the escape key and delay

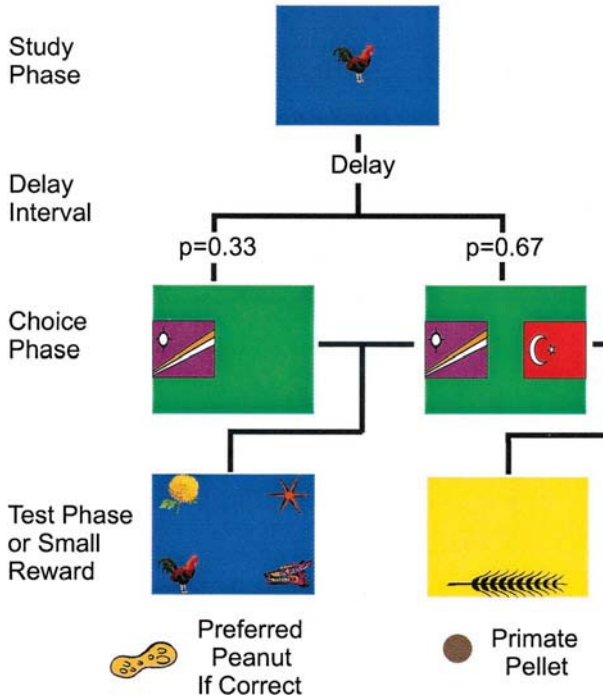


Fig. 1 The task structure in Hampton (2001)

245 duration is also consistent with the monkey having based his decision on the
 246 duration of the delay since the original image was presented. He may simply have
 247 learned that test taking tends to have a happy result when the trial started a short
 248 time ago, and an unhappy result when it started a long time ago. That would be to
 249 base the decision on a first-order representation of the delay rather than a meta-
 250 representation of the memory trace.

251 To test the first-order representational hypothesis against the meta-representa-
 252 tional hypothesis, Hampton compared the accuracy of the monkey's performance on
 253 trials like those described above when he chose to take the test, with trials when he
 254 was forced to take the test. Approximately one-third of trials at each delay duration
 255 were forced trials. In these forced trials, only the test flag appeared at the choice
 256 stage. Hampton found that, when tests were forced, the proportion of trials in which
 257 the monkey correctly identified the sample picture declined as the duration of the
 258 delay increased, but when tests were chosen, accuracy remained high even at longer
 259 delays. If, in choice trials, the monkey's decision to press the test key or the escape
 260 key had been based solely on delay duration, one would have expected to see the
 261 same relationship between accuracy and delay duration in chosen and forced trials.
 262 Given that it declined with increasing delay in forced trials—presumably because
 263 the monkey was, in fact, forgetting the sample—then it should also have declined,
 264 at the same rate, in chosen trials because, by hypothesis, the monkey was using
 265 choice trials only to avoid taking the test at longer delays, and this could not be done in

266 forced trials. The observed pattern—preserved accuracy at longer delays in chosen
 267 trials—implies that the monkey’s use of the escape key enabled him, at *each* of the
 268 longer delays, to avoid taking the test when he was unlikely to pass it. The strength
 269 of the monkey’s memory trace on any given trial will have been a powerful
 270 determinant of whether he was likely to pass the test. Therefore, having excluded
 271 the possibility that the monkey was basing his decisions solely on delay duration,
 272 the accuracy-preserving effect of his decisions makes it plausible that they were
 273 modulated by the strength of the monkey’s own memory of the stimulus.

274 Foote and Crystal (2007) recently reported evidence of metamemory in rats.
 275 Although their procedure is similar to Hampton’s in many respects, there were two
 276 crucial differences: the equivalent of the escape key (a hole into which the rat could
 277 poke its nose) was present, but not illuminated, on forced trials; and the rats made
 278 their discriminative responses at constant locations over trials. The first of these
 279 deviations from Hampton’s procedure means that weaker performance in difficult
 280 forced tests than in difficult chosen tests could have been related to erroneous entry to
 281 the escape hole on forced trials. The second raises the possibility that each rat’s
 282 choice of the test key versus the escape key was regulated by the position of its body
 283 relative to the response keys—an object-level, public motor cue. The potential for
 284 use of public motor cues—stimuli generated by the animals’ own movements, rather
 285 than an internal representation—also makes the results of ‘confidence judgement’
 286 experiments less compelling than Hampton’s. In experiments of this kind (e.g.,
 287 Kornell et al. 2007) commentary key responses are made immediately after the
 288 monkey has made its discriminative response. Therefore, commentary key responses
 289 could be controlled, not by an internal representation, but by some publicly
 290 observable feature (e.g., latency) of the preceding response.

291 It would be good to see a replication of Hampton (2001), in which the same result
 292 was obtained for more than one monkey, but since we don’t think meta-
 293 representation is sufficient for consciousness, further experiments would be needed
 294 in any event to show that this is the type of metamemory that does the trick. We
 295 explain in section “[Testing for meta-memory](#)” below what kind of additional
 296 empirical work should be done to test for consciousness-involving metamemory.

297 **Meta-memory: high level meta-representation**

298 Low level meta-representation is not sufficient for consciousness

299 We saw in the last subsection that Hampton (2001) offers evidence that a rhesus
 300 monkey can solve the memory discrimination task without relying on an external
 301 cue, instead using the strength of its own perceptual memory state as an internal cue
 302 for whether or not it is likely to succeed on a subsequent matching-to-sample task.
 303 We agree with Hampton that that is not yet evidence for consciousness in monkeys.
 304 However, we argue in this subsection that Hampton’s result is indeed evidence for
 305 meta-representation.

306 As Reder (1996) has argued, there are important differences in the literature about
 307 what ‘metacognitive’ amounts to. Perner (1991) identifies meta-representational

308 states in terms of their contents: they are representations of ‘representational
309 relations’. Similarly, we take meta-level contents to be those which concern the
310 thinker’s own representational states (e.g., *I am visually representing a red rose on*
311 *the table*). Object-level states do not (e.g., *there is a red rose on the table*).

312 The results of Hampton’s experiment suggest that the monkey had an internal
313 state, connected to its opt-out behaviour, which covaried with the strength of its
314 recall of the visual stimulus, irrespective of which particular stimulus was being
315 recalled. This accords with Hampton’s suggestion that the monkey might have an
316 internal “flag” for the presence or absence of a memory (p. 5362). Why think
317 the internal “flag” is a representation at all? Shea (2007) argues that meeting the
318 following condition is sufficient to count as a species of internal representation: the
319 animal acquires a new internal state R as a result of learning, the internal state R
320 correlates with the state P of another system and leads to behaviour, and the
321 behavioural output in response to R makes sense in the light of the property P with
322 which it correlates (to state it informally). The monkey’s internal “flag” is an R that
323 meets that sufficient condition. The property with which it correlates happens also to
324 be internal: the presence or strength of a perceptual memory trace.

325 What is the content of this representation? Intuitively, the monkey makes a
326 mistake if it opts for the test when it doesn’t remember the stimulus, and it also
327 makes a mistake when it opts out of the test when it does have an accurate memory
328 trace of the stimulus. These intuitive correctness conditions line up with a plausible
329 account of the function of the representation. Its purpose is to keep track of
330 memories. Without offering a full-scale theory of content, these considerations still
331 suggest that the most plausible content for this internal state is meta-representa-
332 tional—something like *I have a memory of a visual stimulus*.¹

333 We have been working with a sufficient condition for being a representation that
334 is relatively easy to meet. There is no reason to think of the representation as
335 ‘cognitive’ in the way that term is typically used in the metacognition literature.
336 Relatively simple systems like those found in computers, subpersonal brain
337 processing and animal signalling contain states with correctness conditions or
338 satisfaction conditions. They would not count as ‘cognitive’. None displays the kind
339 of psychological sophistication of human beliefs and desires, say (although they
340 may in other respects be more sophisticated). We call them low level represen-
341 tations. We use ‘low level’ versus ‘high level’ not as a value judgement, but to mark
342 this particular kind of variation in psychological sophistication of the representa-
343 tional states. That distinction is orthogonal to the distinction between object-level
344 and meta-level contents. High level representations like beliefs can have both
345 object-level contents (*there was a red rose*) and meta-level contents (*I can*
346 *remember a red rose*). Low level representations may also have meta-level contents
347 (e.g., when a computer keeps track of its own memory registers).

348 As we use the term, low level representations are non-conceptual, need not be in
349 the space of reasons, may be coarse-grained, need not be at the personal level, and
350 do not divide into a range of different mental attitudes such as believing, desiring,

¹ Although we use a structured (linguistic) representation to convey the content, we are not suggesting that the monkey’s representational state has constituent structure or conceptual content.

351 imagining and intending. Carruthers (2007) offers an account of animal perfor-
 352 mance in many metacognition experiments in terms of belief and desires with
 353 merely object-level contents. These explanations nevertheless presuppose a
 354 relatively high level of psychological sophistication because they attribute to
 355 animals states with the functional sophistication of beliefs and desires. The
 356 considerations we have offered suggest that Hampton's monkey may have a meta-
 357 variety of a relatively low level species of representation. Relatively little is needed
 358 for a new state driving rewarded behaviour to count as a meta-representation
 359 according to the sufficient condition mentioned above. Such meta-representations
 360 can arise in systems like current computers which are very unlikely to be conscious.
 361 Accordingly, we argue that this kind of low level meta-representation is not, on its
 362 own, good evidence for consciousness. Something must be added if it is to be turned
 363 into a plausible candidate for our condition *C*. In this subsection we investigate
 364 possible additions, to arrive at a characterisation of a metamemory mechanism
 365 which is plausibly consciousness-involving, which we call 'high level Meta-
 366 Memory', or just 'Meta-Memory'.

367 To turn it into an appropriate condition *C*, we need the mechanism for meta-
 368 representation to meet some further conditions: condition *C* = low level meta-
 369 representation *plus* $X + Y + Z$. As we've said already, some of these conditions
 370 may concern neural structures and processes, for example particular brain areas
 371 (maybe only meta-representation in the prefrontal cortex is consciousness-involv-
 372 ing) or processes (perhaps synchrony at the gamma-wave frequency is required).
 373 Since our aim is a functional characterisation of the consciousness-involving
 374 mechanism, we focus on additional requirements that can be characterised
 375 functionally (the *X* of $X + Y + Z$, as it were).

376 Philosophers have proposed various candidates for this additional factor *X*. Some
 377 argue that the meta-representation must have propositional structure (Rosenthal
 378 2005). But to have thoughts with propositional structure requires a sophisticated
 379 capacity with powers akin to those of linguistic processing, which returns us to the
 380 problem of tying consciousness too tightly to something like language, and so ruling
 381 out animal consciousness at the start of the enquiry. Rosenthal does not claim that
 382 only creatures with language can have thoughts with propositional structure, but
 383 without the capacity for linguistic communication, it is very hard to gather evidence
 384 that an animal does have thoughts with propositional structure. In particular, the
 385 ability to categorise together a range of different stimuli, which is often the basis for
 386 studying 'concepts' in non-linguistic animals, is consistent with such generalisa-
 387 tion behaviour being mediated by propositional structure or by non-conceptual
 388 representations. As a result, a test of consciousness that relies on finding
 389 representations with propositional structure would be hard to apply in non-linguistic
 390 animals. It is possible that the potential for consciousness does, in fact, depend on
 391 the capacity to give a verbal report, or on some important correlate of that ability—for
 392 example, possession of a language of thought, or the potential to code mental
 393 contents in propositionally-structured form. However, the case in favour of
 394 language-dependence and its cousins is not currently so strong that it justifies blank
 395 denial that animals are conscious. If we were to make this assumption, we would not
 396 only risk a major Type 2 error (concluding the phenomenon is absent when it is



397 present), but also miss an opportunity to use the elucidation of conditions for
 398 the investigation of animal consciousness to clarify and extend theories of
 399 consciousness.

400 In searching for the additional factor *X*, other philosophers argue that only fine-
 401 grained contents are made conscious by meta-representation (Carruthers 2000). That
 402 proposal depends upon all perceptual experiences having such fineness of grain
 403 (cf. the experienced location of a touch on your back). But it does attempt to connect
 404 with the kinds of conscious experience that seem, from the first person perspective,
 405 to be involved in memory discrimination tasks, even if their fineness of grain is a
 406 contingent feature with respect to their being conscious. A third proposal is that
 407 meta-memory brings the perceptual memory into the space of reasons (McDowell
 408 1994), making it available for the rational control of action (Baars 1988; Dehaene
 409 and Naccache 2001).² That, too, connects with an intuitive first-person perspective
 410 on memory discrimination tasks. When we use our own conscious recall of the
 411 perceptual stimulus to form an internal prediction of whether or not we will be able
 412 to matching-to-sample, and then act on that prediction, it seems that information
 413 about the stimulus is thereby available to inform any kind of voluntary action (it is
 414 'poised' to be acted on in any of a variety of ways). We are not suggesting that this
 415 would be the only way that human subjects could perform a Hampton-style memory
 416 discrimination task. But we argue that it *is* plausible that, when subjects succeed on
 417 the task by making use of their conscious perceptual memory of the stimulus,
 418 information about that perceptual memory is available to be consumed by any
 419 action system.

420 It is notoriously hard to spell-out this seeming availability. It has been argued that
 421 it is a distinctive functional property of human declarative memory (which is taken
 422 to be conscious) that subjects are able to discern the presence and absence of such
 423 memories (Tulving and Schacter 1990). Global availability is related, but goes
 424 further. It has at least two aspects. First, the way I keep track of whether I remember
 425 the stimulus is not proprietarily connected to a particular external cue, but is
 426 modulated in the same way by quality of the perceptual stimulus, delay since the
 427 stimulus offset, distraction, etc. Secondly, the representation of whether I remember
 428 can be deployed in the control of a range of different actions, rather than being
 429 dedicated to the service of only one project. In short, considering the philosophical
 430 positions and reflecting on conscious memory discrimination tasks from the first-
 431 person perspective brings us to the following characterisation of a potential
 432 additional factor *X*. In performing a memory discrimination task in a way that
 433 depends upon my consciously recalling the perceptual stimulus, I seem to have a
 434 representation, cued by my own memory trace of the stimulus, which is tokened in a
 435 variety of situations, and which is available for the control of a range of different
 436 actions and could be deployed to different ends were I given a different task. Such a
 437 representation would likely count as a meta-representation, following the discussion
 438 above, but it is a meta-representation which, in addition, can be tokened in a variety

2FL01 ² The idea of incorporation in the space of reasons derives from McDowell 1994. The condition we arrive
 2FL02 at below is closer to the global availability for the rational control of action discussed by Baars, and by
 2FL03 Dehaene and Naccache, which is less demanding than McDowell's notion.

439 of different situations and deployed to control a range of different actions. That
 440 requirement, when added to meta-representation, turns it into a plausible candidate
 441 for a functional characterisation of a consciousness-involving mechanism.

442 In sum, our functional characterisation, susceptible to empirical investigation in
 443 other animals, of the mechanism deployed in meta-memory (i.e., high level
 444 consciousness-involving metamemory) is as follows.

445 Condition *C*

446 The subject³ represents that she has a memory of a perceptual stimulus, where
 447 that meta-representation can be tokened in a variety of different situations and
 448 can be deployed to control a range of different actions.⁴

449 We have generated condition *C* by introspective reflection on our own case. That
 450 is only a weak source of evidential support. It is enough, however, to be a plausible
 451 basis for generating a hypothesis for empirical test. It is a substantial empirical issue
 452 whether condition *C* does in fact correlate in humans with conscious recall of a
 453 perceptual stimulus as measured by subjects' verbal reports. That is, do human
 454 subjects meet condition *C* only when they report being conscious, and are they ever
 455 conscious without meeting condition *C*? The argument above makes it plausible, but
 456 does not prove, that condition *C* may correlate with other measures of conscious-
 457 ness—which is enough to motivate a proper empirical investigation.

458 Testing for meta-memory

459 Our condition *C* is susceptible to empirical test. The question to ask is whether
 460 performance on one type of memory discrimination task transfers readily to other
 461 memory discrimination tasks; that is, to use a triangulation approach (Campbell
 462 1954; Heyes 1998). Two categories of transfer test must be combined to show that
 463 an animal meets condition *C* in a Hampton-type memory discrimination task. The
 464 first category of experiments demonstrate decoupling of the metamemory ability
 465 from any particular perceptual cue. For example, does an ability to solve Hampton's
 466 task where memory is degraded by a delay between initial stimulus presentation and
 467 the matching task transfer to trials where the accuracy of memory depends instead
 468 on variations in the duration or the intensity of initial stimulus presentation? The
 469 second category of transfer task looks for output generalisation: the ability to make
 470 use of the metamemory in a range of different tasks. For example, we might ask
 471 whether the type of representation that regulates opt-out behaviour in Hampton's
 472 memory test (matching-to-sample) could also be used to guide behaviour in a

3FL01 ³ We use 'subject' to refer to the organism or system which encounters the perceptual stimulus, has a
 3FL02 memory of it, and tokens a representation of that memory. We do not presuppose that being a subject in
 3FL03 this sense involves a sense of self.

4FL01 ⁴ We are deliberately vague about the modal claim 'can be deployed', since the aim is to match
 4FL02 the intuitive ease or difficulty with which information about conscious representations can be deployed in
 4FL03 the rational control of a range of different actions. For present purposes, we do not need to complete the
 4FL04 separate project of making that notion more precise. The rough idea is that the meta-representation could
 4FL05 be used for new projects simply by the animal changing its preferences or by it moving to an environment
 4FL06 with a different reward structure, without having to undergo further learning in the domain of keeping
 4FL07 track of perceptual recall (i.e., without having to undergo further meta-representational development).

473 different memory test (non-matching-to-sample). It seems very plausible that these
 474 additional tests would be satisfied. If Hampton is right that the monkey's opt-out
 475 behaviour is driven by an internal "flag" tied to the perceptual memory trace, then
 476 all ways of degrading that memory trace (delay, stimulus duration, stimulus
 477 intensity) would have the same effects. But that still needs to be tested. And it
 478 remains an open possibility that the animals perform the experiment in an
 479 informationally-encapsulated way: their training may have allowed them to keep
 480 track of the way the memory trace varies in the given experimental set up, but
 481 without being able to carry that over to situations where the memory trace varies in
 482 other ways, or where the information about the memory trace has to be used for
 483 different actions.

484 Obviously, these transfer experiments would be far from trivial. It would take a
 485 huge amount of work to design and implement effective experiments testing for the
 486 presence or absence of a mechanism meeting condition *C* in even just one other
 487 species. But there is no difficulty in principle with carrying out such investigations.
 488 Our condition *C* is both plausibly consciousness-involving in humans and yet
 489 susceptible to empirical test in other animals. In the remainder of this section we
 490 outline in a little more detail some potential experimental paradigms.

491 *Transfer across perceptual cues*

492 Testing for transfer across situations would be relatively straightforward, and has
 493 been discussed previously in the experimental literature on metacognition in
 494 animals (e.g., Inman and Shettleworth 1999). For example, monkeys would first be
 495 trained on Hampton's task, in which the strength of the animal's memory for the
 496 initial stimulus is manipulated by varying retention interval, i.e., the delay between
 497 presentation of the initial stimulus and the point at which the monkey has the choice
 498 of touching the test flag or the escape flag. Then, once the monkeys were responding
 499 in a way that suggests high level meta-memory—choosing to escape more often on
 500 long than short delay trials—occasional probe trials would be introduced. In these
 501 probe trials, the retention interval would be fixed and of relatively short duration,
 502 but the duration of the initial stimulus would vary. Sometimes it would be very
 503 brief, making the stimulus hard to encode and therefore to remember, and on other
 504 probe trials the initial stimulus would be on the screen for a longer period, making it
 505 easy to encode and remember. If a monkey selected the test flag in a probe trial, he
 506 would proceed to the usual, four-choice matching-to-sample test, but the test
 507 outcome would not be contingent on his response; he would be rewarded (or not
 508 rewarded) regardless of the image he selected. Therefore, and crucially, the
 509 monkeys would not have the opportunity to learn across probe trials that initial
 510 stimulus duration predicts test outcomes. Under these conditions, if meta-memory is
 511 indeed driving the monkeys' opt-out behaviour in the main task, then one would
 512 expect them to opt-out more often in probe trials with short than with long stimulus
 513 durations. This would be expected because, according to the meta-memory
 514 hypothesis, opt-out behaviour in the main task depends on the strength of a memory
 515 trace, not on the duration of the retention interval per se, and therefore the animal's
 516 tendency to take the test when the memory is strong but not when it is weak should



517 persist when memory strength varies with stimulus duration rather than retention
518 interval.

519 As with any single experiment, the outcome of this experiment would not be
520 conclusive. Non-contingent reward on probe trials would ensure that the animals
521 could not learn in the course of the experiment that stimulus duration predicts test
522 outcomes, but it is not impossible that the animals would have learned this from their
523 day-to-day experience before the experiment began. Transfer in a variety of different
524 types of probe test, in which memory strength was manipulated not only by stimulus
525 duration but also by, for example, the presentation of distractors before or after
526 stimulus presentation, would strengthen the case for meta-memory. However, to rule
527 out the possibility that success on probe trials was due to pre-experimental learning
528 about relationships between perceptual cues and test outcomes, it would be necessary
529 to use a novel, and possibly invasive, manipulation. If monkeys showed transfer in
530 probe trials where memory strength was manipulated by direct neurochemical or
531 neuro-electrophysiological means, and if one included appropriate sham controls,
532 then we could be confident that pre-experimental learning was not responsible.

533 *Transfer across actions/outputs*

534 In Hampton's experiment, monkeys touched a flag of one colour to take the four-
535 choice memory test and a flag of different colour to escape the test. To assess
536 transfer across action types, initial training would be followed by the introduction of
537 probe trials with different response requirements. For example, the monkeys might
538 be required to pull one of two levers, rather than to touch one of two flags, to make
539 their choice, or the matching-to-sample test might be changed to a non-matching-to-
540 sample test. In the latter case, two images would appear on the screen, the initial
541 stimulus and an alternative, and the monkey would be rewarded only if he touched
542 the alternative image. Naturally it would take a while for the monkeys to learn the
543 new contingencies—that pulling the left lever activates the test, or that non-
544 matching performance is required in the two-choice test—but if their performance
545 on the main task depends on meta-memory, then eventually they should show the
546 same tendencies in probe trials as in trials on the main task, i.e., to opt-out more
547 often when the retention interval was long, and to show greater accuracy at longer
548 intervals in choice trials than in forced trials.⁵

549 **The payoff**

550 Amongst many tasks that may be consciousness-involving, we have examined
551 metamemory. Hampton (2001) shows that monkeys can predict whether they are
552 themselves likely to succeed at a visual matching-to-sample task. Although not

5FL01 ⁵ Even if monkeys passed all of these transfer tests, it could be argued that their memory state is merely
5FL02 'driving' their choice behaviour; that it plays an important causal role in generating their behaviour, but
5FL03 not by virtue of being understood by the animal *as* a memory. If consciousness of a memory were thought
5FL04 to require understanding the internal state as a memory, then more demanding empirical tests would be
5FL05 required, like those developed in the literature on theory of mind in nonhuman animals.

553 conclusive, his results suggest that one of his monkeys used a meta-representation of
 554 its own perceptual recall, rather than any external cue, to perform this task. However,
 555 meta-representation is not, on its own, plausibly good evidence for consciousness.
 556 What more is needed? By setting ourselves the objective of finding a condition which
 557 is open to empirical confirmation and disconfirmation in animals, we have avoided
 558 anthropocentric answers, and thus conditions which may correlate only in humans
 559 with the presence and absence of consciousness. We labelled the result high level
 560 meta-memory: a subject's representation of her memory of a perceptual stimulus,
 561 where that meta-representation can be tokened in a variety of different situations and
 562 can be deployed to control a range of different actions. Meta-representations which
 563 meet that further condition are plausibly good evidence for consciousness. And first-
 564 person reflection on consciously-performed memory discrimination tasks sug-
 565 gests that our conscious recall of the perceptual stimulus does indeed meet this
 566 condition, although that prima facie case must be substantiated by further empirical
 567 investigation. Thus, the payoff from our investigation of animal consciousness is not
 568 just to show, in the face of methodological scepticism, that it is an empirically-
 569 tractable question. It has also led us to a sharper conception of the nature of
 570 consciousness itself, in humans and other animals, forcing us to specify in greater
 571 detail the functional profile of the mechanisms deployed by subjects when they rely
 572 on consciousness to solve a memory discrimination task.

573 Those who view higher order thought as necessary for consciousness can take our
 574 condition *C* as a candidate for upgrading meta-representation into a sufficient
 575 condition for consciousness. But taking condition *C* as partly constitutive of
 576 consciousness would join higher order thought theories in making consciousness a
 577 matter of having certain dispositions. Our claim is less controversial: that
 578 discovering meta-memory in animals would be good evidence that they are
 579 conscious. That is a substantial claim, of considerable interest whether or not higher
 580 order theories are right. But this story has a final twist. We have been assuming
 581 throughout that, to solve a memory discrimination task without using an external
 582 cue, a subject would have to use some additional internal state, over and above its
 583 perceptual memory. We argued that, if so, the new representation would likely be a
 584 meta-representation, rather than having object-level contents. Our concern was to
 585 see what needed to be added to meta-representation, to turn it into a plausibly
 586 consciousness-involving mechanism. However, once we've seen that additional
 587 factor *X*, we can ask whether meta-representation is a necessary part of the
 588 evidential condition *C*, or whether the factor *X* would, on its own, be good evidence
 589 for consciousness. Assessing the theoretical considerations in favour of that
 590 hypothesis and making suggestions for testing it empirically (would be a paper in its
 591 own right. We restrict ourselves to observing that our factor *X* is similar to
 592 Dehaene's global workspace hypothesis (Dehaene and Naccache 2001), which is
 593 formulated as a necessary and sufficient condition for a mechanism to be conscious.

594 **Acknowledgments** For helpful discussion and comments, the authors would like to thank Tim Bayne,
 595 Alan Cowey, Martin Davies, David Shanks, Kim Sterelny and an anonymous referee for this journal.
 596 NS's research is supported by the OUP John Fell Research Fund, the James Martin twenty-first Century
 597 School, the Oxford Centre for Neuroethics and the Mary Somerville Junior Research Fellowship,
 598 Somerville College.

599 **Open Access** This article is distributed under the terms of the Creative Commons Attribution Non-
 600 commercial License which permits any noncommercial use, distribution, and reproduction in any med-
 601 ium, provided the original author(s) and source are credited.

603 References

- 604 Baars B (1988) A cognitive theory of consciousness. Cambridge University Press, Cambridge
 605 Campbell DT (1954) Operational delineation of “what is learned” via the transposition experiment.
 606 Psychol Rev 61(3):167–174
 607 Carruthers P (2000) Phenomenal consciousness: a naturalistic theory. Cambridge University Press,
 608 Cambridge
 609 Carruthers P (2007) Meta-cognition in animals: a skeptical look. *Mind Lang* 22
 610 Clark RE, Squire LR (1998) Classical conditioning and brain systems: the role of awareness. *Science*
 611 280:77–81
 612 Clark RE, Manns JR, Squire LR (2001) Trace and delay eyeblink conditioning: contrasting phenomena of
 613 declarative and nondeclarative memory. *Psychol Sci* 12:304–308
 614 Cowey A, Stoerig P (1995) Blindsight in monkeys. *Nature* 373(6511):247–249
 615 Cowey A, Stoerig P (1997) Visual detection in monkeys with blindsight. *Neuropsychologia* 35(7):929–939
 616 Dehaene S, Naccache L (2001) Towards a cognitive neuroscience of consciousness: basic evidence and a
 617 workspace framework. *Cognition* 79(1–2):1–37
 618 Foote AL, Crystal JD (2007) Metacognition in the rat. *Curr Biol* 17:551–555
 619 Hampton RR (2001) Rhesus monkeys know when they remember. *Proc Natl Acad Sci USA* 98(9):5359–
 620 5362
 621 Hampton RR, Zivin A, Murray EA (2004) Rhesus monkeys discriminate between knowing and not
 622 knowing and collect information as needed before acting. *Anim Cogn* 7:239–246
 623 Heyes CM (1998) Theory of mind in nonhuman primates. *Behav Brain Sci* 21(1):101–114 (discussion
 624 115–148)
 625 Heyes CM (2008) Beast machines? Questions of animal consciousness. In: Davies M, Weiskrantz L (eds)
 626 *Frontiers of consciousness*. OUP, Oxford
 627 Inman A, Shettleworth SJ (1999) Detecting metamemory in nonverbal subjects: a test with pigeons. *J Exp*
 628 *Psychol Anim Behav Process* 25:389–395
 629 Jack AI, Shallice T (2001) Introspective physicalism as an approach to the science of consciousness.
 630 *Cognition* 79(1–2):161–196
 631 Kornell N, Son LK, Terrace HS (2007) Transfer of metacognitive skills and hint seeking in monkeys.
 632 *Psychol Sci* 18(1):64–71
 633 McDowell JH (1994) *Mind and world*. Harvard University Press, Cambridge
 634 Perner J (1991) *Understanding the representational mind*. MIT Press, Cambridge
 635 Perruchet P (1985) A pitfall for the expectancy theory of human eyelid conditioning. *Pavlov J Biol Sci*
 636 20:163–170
 637 Perruchet P, Cleeremans A, Destrebecqz A (2006) Dissociating the effects of automatic activation and
 638 explicit expectancy on reaction times in a simple associative learning task. *J Exp Psychol Learn*
 639 *Mem Cogn* 32(5):955–965
 640 Reder LM (1996) Different research programs on metacognition: are the boundaries imaginary? *Learn*
 641 *Individ Differ* 8(4):383–390
 642 Rosenthal D (2005) *Consciousness and mind*. OUP, Oxford
 643 Shea NJ (2007) Consumers need information: supplementing teleosemantics with an input condition.
 644 *Philos Phenomenol Res* 75(2):404–435
 645 Smith DJ, Shields WE, Washburn DA (2003) The comparative psychology of uncertainty monitoring and
 646 metacognition. *Behav Brain Sci* 26:317–373
 647 Stoerig P, Zontanou A, Cowey A (2002) Aware or unaware: assessment of cortical blindness in four men
 648 and a monkey. *Cereb Cortex* 12(6):565–574
 649 Tulving E (2005) Episodic memory and autoeosis: uniquely human? In: Terrace HS, Metcalfe J (eds)
 650 *The missing link in cognition*. OUP, Oxford, pp 3–55
 651 Tulving E, Schacter DL (1990) Priming and human memory systems. *Science* 247:301–306
 652 Weiskrantz L (1986) *Blindsight: a case study and implications*. OUP-Clarendon Press, Oxford
 653 Weiskrantz L (1995) The problem of animal consciousness in relation to neuropsychology. *Behav Brain*
 654 *Res* 71:171–175
 655