Acta Psychologica 164 (2016) 56–64

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

The Doubting System 1: Evidence for automatic substitution sensitivity

Eric D. Johnson [a,*], Elisabet Tubau [a], Wim De Neys [b]

[a] Department of Basic Psychology and IR3C, University of Barcelona, Barcelona, Spain
[b] CNRS, LaPsyDE (CNRS Unit 8042), Paris Descartes University, Paris, France

ARTICLE INFO

Article history:
Received 19 May 2015
Received in revised form 2 December 2015
Accepted 14 December 2015
Available online 23 December 2015

Keywords:
Cognitive reflection
Bat-and-ball problem
Dual process
Executive resources
Reasoning
Decision making
Bias

ABSTRACT

A long prevailing view of human reasoning suggests severe limits on our ability to adhere to simple logical or mathematical prescriptions. A key position assumes these failures arise from insufficient monitoring of rapidly produced intuitions. These faulty intuitions are thought to arise from a proposed substitution process, by which reasoners unknowingly interpret more difficult questions as easier ones. Recent work, however, suggests that reasoners are not blind to this substitution process, but in fact detect that their erroneous responses are not warranted. Using the popular bat-and-ball problem, we investigated whether this substitution sensitivity arises out of an automatic System 1 process or whether it depends on the operation of an executive resource demanding System 2 process. Results showed that accuracy on the bat-and-ball problem clearly declined under cognitive load. However, both reduced response confidence and increased response latencies indicated that biased reasoners remained sensitive to their faulty responses under load. Results suggest that a crucial substitution monitoring process is not only successfully engaged, but that it automatically operates as an autonomous System 1 process. By signaling its doubt along with a biased intuition, it appears System 1 is "smarter" than traditionally assumed.

1. Introduction

In the face of difficulty, human reasoners often appear to forego the effortful processing that may be required and opt instead for less demanding intuitive responses (Kahneman, 2011). While many fast and frugal heuristics are no doubt adaptive in complex and reoccurring environments (Gigerenzer, 2007), thinking fast can also lead to quite embarrassingly erroneous responses in less routine settings. Quickly consider the following example:

> A bat and a ball together cost $1.10.
> The bat costs $1 more than the ball.
> How much does the ball cost?

Intuitively, the answer "10 cents" quickly springs to mind. In fact, a majority of university students, including those from elite schools such as MIT and Harvard, respond with this *intuitive*—but *incorrect*—answer (e.g. Bourgeois-Gironde & Van der Henst, 2009; Frederick, 2005). If a bat costs $1 more than a 10-cent ball, the bat itself must cost $1.10. Summing up, a $1.10 bat + a $0.10 ball would equal $1.20, not $1.10 as stated in the problem. Does this imply that highly educated young adults think that '110 + 10' = '110'? Of course not. Rather, it suggests that even educated reasoners often do not invest the necessary effort needed to correct their initial intuitions, and instead settle for a quickly derived response.

This resistance to cognitive expenditure, or "miserly" thinking, has been most famously characterized by Kahneman (2011), Kahneman and Frederick (2002). According to this dual-process view, when people are confronted with a difficult question, an autonomous System 1 quickly and unconsciously substitutes an easier question in its place. In the bat-and-ball problem, this presumably involves the swapping of the critical relational "more than" statement with a simpler absolute interpretation. That is, people will read "the bat costs more than" as simply "the bat costs", and therefore perhaps ironically give the right answer to the wrong question. Correcting this faulty intuition is assumed to depend on the active monitoring of System 1 by a deliberate and resource-demanding System 2. Due to the human tendency toward miserly or "lazy" thinking, however, this monitoring process typically fails to engage. Without the engagement of System 2, we blindly go with the substituted System 1 response.

More recent work, however, has questioned the extent to which this substitution process goes unnoticed. De Neys, Rossi, and Houdé (2013) solicited participants' judgments of confidence in their response after solving the standard bat-and-ball problem or the following control version:

> A magazine and a banana together cost $2.90.
> The magazine costs $2.
> How much does the banana cost?

In this control version, people will tend to parse the $2.90 into $2 and 90 cents just as naturally as they parse $1.10 in the standard version. However, the control version no longer contains the relative statement

* Corresponding author at: Departament de Psicologia Bàsica, Facultat de Psicologia, Universitat de Barcelona, Passeig de la Vall d'Hebron, 171, 08035 Barcelona, Spain.
E-mail address: eric.johnson@ub.edu (E.D. Johnson).

http://dx.doi.org/10.1016/j.actpsy.2015.12.008
0001-6918/

("$2 more than the banana") which triggers the substitution. That is, the control version directly presents the easier statement that participants are supposed to be unconsciously substituting in the standard version. If participants are completely unaware that they are substituting when solving the standard version, the standard and control version should be isomorphic and response confidence should not differ. De Neys et al. (2013) observed, however, that participants were much less confident when they erroneously substituted the "10 cents" response on the standard bat-and-ball problem compared to their confidence on the control version (see also Gangemi, Bourgeois-Gironde, & Mancini, 2015, for similar findings). This work suggests that, at least at some level, we are not blind to the substitution process—even biased reasoners showed elementary substitution sensitivity. If this is true, however, it raises an even more fundamental question regarding the source of this sensitivity.

In the present study we contrast two possible origins of this previously observed substitution sensitivity. First, this detection process may be part of a monitoring component of System 2, as suggested by Kahneman (2011), Kahneman and Frederick (2002). On this view, although a supervisory System 2 may not be allocating sufficient resources to the override processes needed to solve the bat-and-ball problem, it is to some extent monitoring for inappropriate output. Bluntly put, System 2 would be more active than typically assumed. However, a second possibility is that this substitution sensitivity arises out of an autonomous System 1 process. On this account, System 1 does not ignorantly throw out an answer whose outcome is at the complete mercy of a vigilant, interventionist System 2. Rather, it sends with its rapid approximation a signal of doubt. Simply put, while System 1 may not be "intelligent" in the traditional sense, neither is it as "dumb" or blind as characteristically assumed.

These two possibilities can be teased apart using the basic processing assumptions of dual process theories. System 1 processes are thought to operate automatically, out of the grip of more controlled, demanding System 2 processing which depends on the availability of executive resources (Evans, 2008; Evans & Stanovich, 2013). The locus of substitution sensitivity can therefore be tested by experimentally manipulating the executive load placed on participants as they reason with the bat-and-ball problem. If detecting an erroneous substitution process is in the domain of a deliberate System 2, then under a resource-demanding load reasoners should not detect this substitution, or this sensitivity should be greatly reduced. If, on the other hand, substitution sensitivity is the work of an automatic System 1 process, then this detection mechanism should be unaffected by load.

In the present investigation we probe this substitution sensitivity in the bat-and-ball problem (and a control version) while reasoning under cognitive load. Four load conditions were used—no load, low load, high load, and extra-high load—to examine the relative contributions of executive resources both for correctly solving the problem and for detecting the presumed substitution when answering with an erroneously substituted response.

In order to validate these findings, we included three different substitution sensitivity measures: Confidence judgments, confidence latencies, and reasoning latencies. Note that the sensitivity findings of De Neys et al. (2013) were based purely on a confidence measure. However, studies investigating basic cognitive control processes in reasoning have shown that decision uncertainty associated with conflict also affects response latencies (Scherbaum, Dshemuchadse, Fischer, & Goschke, 2010; see also Bonner & Newell, 2010; De Neys & Glumicic, 2008; Mevel et al., 2014; Pennycook, Fugelsang, & Koehler, 2012; Stupple & Ball, 2008; Stupple, Ball, & Ellis, 2013; Thompson, Striemer, Reikoff, Gunter, & Campbell, 2003; Villejoubert, 2009). Therefore, if sensitivity arises out of the substitution process then, in addition to reduced response confidence, we should also expect to see longer response times as reasoners attempt to solve the standard version of the task. That is, if reasoners are questioning whether their substituted response is warranted, this uncertainty should translate into increased processing time on the standard bat-and-ball problem relative to a situation where there is no questioning of the immediate intuition (i.e., the control version). Furthermore, latencies for the confidence judgment itself might be affected. If one feels unsure of their response, it may take more time to translate this feeling into a precise estimate of confidence compared to when one is fully confident. Hence, measuring the time it takes to provide a judgment of confidence may provide an additional index of substitution sensitivity.

In sum, if reasoners are sensitive to the substitution process then one can predict that, in addition to previously observed lower confidence ratings, responding to the problem and providing a subsequent judgment of confidence should take longer for standard versus control versions of the task. The key question, however, is whether or not these three detection measures still indicate substitution sensitivity under cognitive load. If this sensitivity depends on the operation of an executive resource-demanding System 2, then its effectiveness should decline under load. However, if substitution sensitivity arises out of autonomous System 1 processes, these measures should be unaffected by load.

2. Experiment

2.1. Method & material

2.1.1. Participants

A total of 324 undergraduate students from the University of Barcelona were recruited for this task in exchange for course credit. Eleven of these students reported being previously familiar with the bat-and-ball problem, and therefore only data from the remaining 313 participants (266 female, 47 male; mean age = 20.50, $SE = 0.28$) was analyzed and reported here.

2.1.2. Reasoning task

The reasoning tasks included a standard and a control version of the bat-and-ball problem introduced above. As in previous work (De Neys et al., 2013), different contextual and numerical contents were used (see Appendix A). One problem presented a bat and ball, the other presented a magazine and banana. In one problem the total cost was $1.10 with one item costing $1 more than the other; in the other problem the total cost was $2.90 with one item costing $2 more than the other. Item contents and values for the standard and control versions were fully counterbalanced across participants, which helps to ensure that any observed effects are general and not driven by the specific material used (e.g. the ease of partitioning 10 from 1.10, or background beliefs about the price of specific items).[1] A blank box with the label "cents" appeared on screen following the problem. Participants therefore typed only their numerical response into the box.

2.1.3. Confidence measure

Immediately following response to either the standard or control version of the reasoning task, participants were asked to indicate how confident they were that their response was correct. Confidence judgments were indicated with a numerical value between 0% (*not at all confident*) and 100% (*completely confident*). As in previous studies (e.g. De Neys, Cromheeke, & Osman, 2011; De Neys et al., 2013), the interest is in the *relative* difference between confidence judgments on the standard substitution version and the control problem. There are numerous reasons for individual variation in absolute ratings of confidence, and a variety of measurement biases may influence the particular value that participants report (e.g. Berk, 2006; Shynkaruk & Thompson, 2006). Accordingly, absolute confidence levels must be interpreted with caution. At the same time, however, it can be assumed that any general bias in the response scale should affect confidence ratings in both standard and control versions. Observing relatively lower confidence

[1] None of these factors had any impact on performance.