In the target paper, I outlined several methodological issues associated with attempts to document category specific deficits; and a potential solution based around certain minimal criteria. The main argument being that an accurate interpretation of patient performance requires a comparison with a normal control group and that the group performs below ceiling. Neither of these requirements is new or one would imagine, especially contentious. It is therefore surprising that a review of the category specific literature reveals no single study that meets these criteria (Laws, in press). Moreover, this has to be viewed in the context that the commonly used analyses (within-patient \( \chi^2 \) and between-subject comparisons with controls at ceiling) produce false positive, false negative and even paradoxical deficits.

The target paper commentaries may be roughly classified into those on the one hand in favour of using controls who are not at ceiling (Sartori and Lombardi; McMullen and Filliter; Rosazza et al.,) and that analyses of healthy controls are critical in themselves (Låg; Marques), through to those who feel that their data already meets the criteria in the target paper (Capitani and Laiacona; Laiacona and Barbarotto; Bright et al.,) or argue that controls are not necessarily required (Marshall and Gurd’s extreme cases approach). This reply will focus primarily on general points of contention raised by the contributors.

Solutions to the Ceiling Effect in Controls

Capitani and Laiacona discuss various possible solutions to address ceiling effects in controls, including: degrading the stimuli in some way to make them more difficult for controls, using controls who perform below ceiling (for example, older subjects) or constructing a more difficult stimulus set.

Both Capitani and Laiacona and Marshall and Gurd reject the stimulus degradation idea because it would introduce unknown variables (see Låg) as also remarked upon in the target article (see comments on Turnbull and Laws, 2000). Although stimulus degradation is a useful tool for examining the factors affecting the category differences in healthy subjects (Låg; Marques; Låg, in press; Laws and Neve 1999; Laws et al., 2002; Laws, 2000), it cannot provide control baselines. This becomes apparent when one considers that different presentation conditions may interact with category (see Låg, in press; Gerlach, 2001; Laws and Neve, 1999).

The solution preferred by Capitani and Laiacona is a “slightly different” group of normal subjects who in this case, consist of 60 elderly subjects (30 males and 30 females aged 71). These elderly subjects used in all of the studies by Capitani, Laiacona and colleagues do however have naming at ceiling with 91.83 ± 12.8% and 96.38 ± 7.33% of the controls naming all of the living things and nonliving things correctly

ii More critically, while the regression method advocated by Capitani and colleagues could potentially address some of the issues raised in the target paper, it is not immune to ceiling effects. In the original article, I mistook the term ‘Name Agreement’ used by Capitani, Laiacona and colleagues’ papers to mean what it typically means, i.e. the number of alternative names produced rather than an indicator that control data were included in the regression.

Normal Individual Variability

As noted by Marques, natural asymmetries of category processing exist in healthy subjects and so, control data are essential to “...even establish
that a real deficit is present”. Therefore, as Sartori and Lombardi remark, we cannot rely upon generic control groups because category performance may be affected by sex, education and age differences. In the context of normal individual variability, Laiacona and Barbarotto refer to several studies documenting an interaction between sex and category (e.g. Capitani et al., 1999; Barbarotto et al., 2002; Laws, 1999, 2003) with males performing better with nonliving things and females with living thingsiii. Findings relating to such individual differences are important both as an aid to understanding category effects and in determining the nature of appropriate controls i.e. they indicate that controls should be sex matched.

Illusions of Double Dissociations beyond Naming

An altogether different issue raised by Laiacona and Barbarotto and by Bright et al. concerns the emphasis on picture naming in the target article. While picture naming does represent just one test of category specificity, it was chosen for a variety of reasons, including the fact that it is the one test common to almost all studies and is used (implicitly or otherwise) as the standard means to documenting category effects. Of course, most studies use many other tests; however, the choice of any other single test is likely to be less informative and less widely used (e.g. drawing, naming-to-description, sentence verification).

If one accepts the methodological problems relating to controls in picture naming studies, it would be unsurprising to see that the same problems extend to other tests in the same studies (or even beyond category specificity as alluded to by Marshall and Gurd). Indeed, a cursory glance at category specific studies (or the wider literature) appears to confirm this suspicion. As an example, let us examine the semantic question data for PL and MF (Laiacona and Barbarotto)iv as a potential double dissociation across living and nonliving things. The absolute scores and the cross-over figure do look convincing, and are typical of the evidence presented for double dissociations (and ultimately the fractionation of cognition). Nevertheless, the normative data for this task come from the same 60 elderly individuals and is again at ceiling (96.7% and 98.6% of subjects answered every living and nonliving thing question correctly). Therefore, the ceiling effect problem with naming also extends to other tests in category specific investigations. At a more abstract level, consider the typical double dissociation in this area of enquiry (which is unreferenced to controls) below and how the interpretation might change when control data are added. This example shows how absolute performance may be misleading. An apparent and typical double dissociation, when referenced to controls, may reveal the same category deficit in both patients (here for living things).

---

iii Marshall and Gurd provide an example in which it seems John Marshall knows nonliving things such as demijohn and tantalus, but not armadillo and platypus. One possible explanation is the interaction between category and sex.

iv Capitani and colleagues refer to several of their cases that “escaped my attention” (re: Barbarotto et al., 1995 [MF]; Albanese et al., 2000 [GR and PL]); Laiacona et al., 2003 [EA]). Three of these cases were actually referenced in the target paper – each being reported on more than one occasion elsewhere: PL (Laiacona and Capitani, 2001); GR (Laiacona et al., 1993); EA (Barbarotto et al., 1996; Laiacona et al., 1997).
Extreme Cases

Marshall and Gurd refer to the method of ‘extreme cases’, where patients, but not normal subjects for example “…make semantic errors when reading individual words…(or) draw clockfaces that only include the numbers 12, 1, 2, 3, 4, 5 and 6”. These examples, however, are dealing with qualitatively different human behaviours. While in category specificity (and many other neuropsychological deficits), the focus is on quantitative differences and it is not known a priori whether normal subjects would show a living or nonliving advantage. In this sense, while extreme cases may be useful, they are unusual and in the minority of published neuropsychological case studies.

If Normal Subjects show a Living Advantage, then why are there so many Living Deficits?

Bright et al. suggest that some of the points I raised in the target article are “ill-judged and contradictory”. The focus of the argument by Bright and colleagues appears to rest on two assumptions that they take to be contradictory: (a) the finding that normals show better naming of living than nonliving things; and (b) the far greater incidence of living deficits reported in patients.

Turning to the first assumption, if materials are matched across category on the usual nuisance variables, controls do show better naming of living than nonliving things (see Laws, 1999, 2000, 2002a, 2003; Laws et al., 2002b; Laws and Neve, 1999). Despite the assumptions that some researchers build into their models of category specificity, I am not aware of any study of normal subjects reporting a naming advantage for nonliving things on matched stimuli. The second assumption is one that I have directly questioned (Laws et al., in press; Laws, in press). Indeed, Laws et al. (in press) have shown that over-reliance on stimuli that engender ceiling effects in controls may inflate the proportion of living cases and underestimate the number of nonliving ones (Laws et al., in press; Laws et al., 2003).

Bright et al. suggest that that I have been over-dissmissive of their data for JBR (Bunn et al., 1998) and RC (Moss et al., 1998). In both cases, the same control data were taken from a colour photo set in 40 young controls and 8 elderly controls. Contrary to what Bright et al. claim, both control groups were at ceiling (living and nonliving things, young controls score 93 vs. 96% correct; and for the 8 elderly controls score 95 vs. 95% correct).

Finally, they refer to the contradictory findings from the case of SE (originally described by Laws et al. (1995), but later also described by Moss et al., 1998; see also Laws, 1998). If a patient produces contradictory results across groups, then it does seem grounds for doubting the findings of both studies and a warning about comparing across different paradigms and materials (see also commentary by Låg; see also Laws, 1998). Second, and critically, they argue that they “…did meet all of Laws’ criteria for comparison to control groups” (my italics), but this is not the case. Three sets of naming data are presented: the full Snodgrass and Vanderwart corpus (using 12 controls) of stimuli unmatched across category; a subset of those stimuli matched for frequency and familiarity, but with no controls; and the colour photos from Bunn et al.; here used within patient comparison and no reference to controls. Hence, none of their data meets the criteria outlined in the target article.

Leaving aside Laws et al. (in press) and Rosazza et al., which demonstrate the approach advocated, and despite the claims made in some commentaries, no study has documented a category effect that meets the criteria outlined in the target article. In other words, none have demonstrated a category effect using: (a) a matched (sex, age etc) control group who; (b) do not perform at ceiling on: (c) stimuli that have the usual cross-category confounding differences controlled; and (d) in which the sample statistics are treated as statistics rather than as population parameters. Moreover, this means that, currently, no soundly established double dissociation between living and nonliving things exists. As noted in the target paper, this does not mean that category effects do not exist. Indeed, the approach outlined above represents a set of minimal criteria that will enhance the accuracy of their documentation and interpretation. While it remains to be determined whether these control problems apply more generally in cognitive neuropsychology (pace Marshall and Gurd), they do extend beyond picture naming and pervasively affect category specific studies.

REFERENCES


---

viii See Crawford and Garthwaite (2002) for a detailed description of the statistical issues and Laws, Gale, Leeson and Crawford (in press) for examples of the use of these statistics to category deficits in Alzheimer’s patients.


