ELSEVIER

Neurobiology of Aging 31 (2010) 2172–2175

NEUROBIOLOGY OF AGING

www.elsevier.com/locate/neuaging

Authors' response to open peer commentary

Response to open peer commentary on the reporting of spurious associations: a reply to "Relating hippocampal neurogenesis to behavior: the danger of ignoring confounding variables" by Dr. Stanley Lazic

Benjamin K. Yee[a,*], Joram Feldon[a,*], Irene Knuesel[a,b]

[a] *Laboratory of Behavioural Neurobiology, Swiss Federal Institute of Technology, Zurich, Switzerland*

[b] *Present address: Institute of Pharmacology and Toxicology, University of Zurich, Zurich, Switzerland*

Received 2 June 2010; received in revised form 29 June 2010; accepted 10 July 2010

Abstract

Often factors related to hippocampal neurogenesis also result in a myriad of confounding changes that might explain or mediate the concomitant effects in memory and learning performance. Dr. Lazic's article (2010) reiterates such concerns in interpreting the biological links between neurogenesis and learning in the context of aging as articulated previously by Baxter and Gallagher (2006). The correlative analysis published by Nyffeler et al. (2008) illustrates the problematic inherent to such an interpretation. We offer here a complimentary approach that is both intuitive and practical in the re-examination of the previously reported data set, which further supports Nyffeler et al.'s (2008) key findings and conclusion.

Keywords: Aging; Correlation; Doublecortin; Nestin; Neurogenesis; Partial correlation statistics

Dr. Lazic's critique focuses essentially on a scatter plot included in a single composite figure appearing in our previous paper (Fig. 6F of Nyffeler et al., 2008) illustrating the relationship between the number of doublecortin (DCX)-positive cells in the hippocampus and performance in a water maze learning task across four groups of rats. The four groups were differentiated by age: 3, 6, and 24 months, with the last group further subdivided into *good* versus *bad* learners based on individual water maze performance. In spite of the presence of a statistically significant linear association between the two concomitant effects of aging, Nyffeler et al. (2008) concluded that the data did not lend support for a link between neurogenesis and memory performance. The total number of new hippocampal neurons as such could not explain the performance difference existed between *good* and *bad* aged subjects. This null effect strengthens the major finding by Nyffeler et al. (2008) concerning the significance of the ratio between nestin-to DCX-immunoreactive cells in the hippocampus. This ratio was markedly shifted from *good* (30%) to *bad* (70%) learners in the aged groups (see Fig. 8 of Nyffeler et al., 2008), and represents a novel finding in the literature, highlighting the importance in distinguishing between different ontological phases of neurogenesis rather than merely focusing on the numbers of new neurons.

To plot or not to plot

Neither the experimental design nor the conclusion reported in Nyffeler et al. (2008) depends on the illustration of the scatter plot under scrutiny. Reporting separately the

* Corresponding authors at: Swiss Federal Institute of Technology, Laboratory of Behavioural Neurobiology, Schorenstrasse 16, C6 CH 8603 Schwerzenbach, Switzerland. Tel.: +044 655 7448; fax: +41 44 655 72 03.

E-mail address: byee@ethz.ch (B.K. Yee); feldon@behav.biol.ethz.ch (J. Feldon).

0197-4580/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.neurobiolaging.2010.07.019

presence of (1) age-dependent cognitive decline, and (2) age-dependent reduction in the total number of DCX-ir neurons in dentate gyrus is already sufficient as validation of the background to our study's aim, which sought to test if animals belonging to the same senescent age group (24 months) but demarcated by cognitive performance could be distinguished by markers of neurogenesis or associated cellular events. This hypothesis essentially requires a statistical comparison of the relevant markers between the *good* and *bad* performers belonging to the same senescent age group.

However, the scatter plot serves as an effective visual aid to illustrate the covariation between the number of DCX-ir cells and performance index across groups, and the Pearson's product moment correlation helps to indicate the statistical strength of this relationship ($R^2 = 0.46$) which by itself does not imply casual relationship. Furthermore, the group identity of individual data points was clearly marked so that it is immediately obvious that the apparent correlation is largely due to the concomitant monotonic effects of aging on DCX-ir cell numbers (represented in the x-axis) and cognitive performance (represented in the y-axis). The correlation is spurious in statistical terms, but it is a robust and meaningful biological observation.

"Reichenbach (1971) proposed that a robust correlation between variables is spurious [acausal] when there is a factor that 'screens off' the correlation and serves as a common cause of the associated variables" (quoted from plato.stanford.edu/entries/paradox-simpson). The common "cause" is readily recognized here as the animals' age. However, immensely profound correlations can often be statistically spurious, such as that between recession speed and distance of galaxies (Hubble's Law), and the tight correlation between black hole mass and velocity dispersion sigma of the stellar bulge in galaxies (the M-σ Relation). Hence, one must not be confused by the dictionary meaning of the term "spurious", which carries a negative connotation implying falsehood. Yet, reporting such correlations requires caution.

Reporting and plotting spurious correlations

Bizon and Gallagher (2003) have provided an example of statistically controlling for the impact of age on their correlative analysis between neurogenesis and memory performance. Individual memory scores were normalized by the respective group's average from each individual score, so that any age difference was eliminated (see their Fig. 3B). This unique approach might have been chosen because these authors were reluctant to treat the grouping variable age as a continuous variable. Lazic (2008, 2010) suggested that such concern is not necessary and we employ here the standard partial correlation approach to reanalyze our data (Chen and Popvich, 2002). This approach, we think, is more intuitive (easily linked conceptually to the original "uncontrolled" correlation) than Dr. Lazic's approach. Fig. 1 illustrates the use of partial correlation plots (Velleman and Welsch, 1981), allowing a visual comparison with the original scatter plot. This is especially useful in such a 3-factor situation, but can be readily extended to higher multifactorial situation.

The statistical conclusions are consistent with Dr. Lazic's reanalysis. The correlation in question dropped from $r = +0.679$ to $r = +0.018$, which matches to the similarly weak correlation ($r = 0.03$) reported by Bizon and Gallagher (2003) using their nonstandard partial correlation. We illustrated here also the other two possible partial correlations: (age → DCX) and (age → memory performance) (with the remaining third factor partial-out). These essentially correspond to what Nyffeler et al. (2008) originally reported separately using analysis of variance (ANOVA); and their persistence when the third factor is partial-out is consistent with the near complete lack of association between "DCX" and "memory performance" when "age" is partial-out.

Nyffeler et al. (2008) reported a highly consistent finding of a decline in hippocampal cell proliferation that is primarily attributed to age, instead of being predictive of cognitive performance in the aged subjects. The reanalyses provided here and by Dr. Lazic reaffirm Nyffeler et al.'s original statement that "it is apparent from the scatter plot that the age-dependent reduction of newborn cells alone did not discriminate between subjects from the 2 subsets of aged rats differing in cognitive abilities" (pp. 12).

Pondering causality

The reference to the Simpson's paradox (plato.stanford.edu/entries/paradox-simpson) by Dr. Lazic is highly relevant to data interpretation not only in the explicit use of correlative analysis, but also in experiments involving treatment factors with independent variables such as drugs, lesions, or genetic mutations. The doctrine that "correlation does not imply causation" arguably also holds in all such cases, because the treatment factors also induce multiple confounding changes that might "screen off" the supposedly causal link between independent and dependent variables. The Simpson's paradox illustrates how a causal stance could become perplexing, even though the arithmetical structure behind is not in itself paradoxical. However, the identification of the confounding mediating factors of some such spurious correlations can shed light on the mechanism involved.

The confounding factor of focus here has been age, which is of course central to the neurobiology of aging. Senescence is a biological phenomenon of interest not simply because of the passage of time (arguably its ultimate cause), but also the associated multitude of biologically meaningful effects: scruffy fur coat or poor skin complexion, cognitive deterioration, or reduced capacity to generate new neurons. Our objective must be to identify the mechanistic structures by which these changes come about in