A recent paper suggested that alcohol consumption might deal a blow to a very sacred cow – scientific productiv-
ity (Grim 2008). Great numbers of comments followed. As noted by Sheil et al. (2008), the “Grim’s study cleverly
highlights a more general lesson. His analyses and presentation follow accepted practices in evolutionary ecology. These are too seldom challenged. It appears we only notice failings when we are motivated by finding ourselves in the
study population.” [italics are mine in all cited text]. In-
deed, this fulfils one of the aims of the Oikos ‘beer study’
– to serve as a mirror for ecological research. In this, the
beer study parallels the famous Sokal’s ‘experiment with
cultural studies’ paper that was a mirror to some areas of
social studies and humanities (Sokal 2008). Elsewhere, I
will provide detailed comments on issues that were dis-
cussed by the critics of the Oikos study. Here, I reply
specifically to two comments published by Web Ecology. Both Moya-Laraño (2008) and Sheil et al. (2008) draw
attention to some issues of general interest. At the same
time, their papers provide examples of some weaknesses of
ecological papers and fallacious arguments so commonly
employed by laymen and scientists. Although I intended
the Oikos study to be fun, my message here is serious. Fo-
cusing on the drawbacks is a crucial condition for progress
in science.

Survey bias: the problems of self-reported data

‘Self-reported’ data might not be reliable (Sheil et al.
2008). First, Sheil et al. refer to a study based on “alco-
hol offenders in a rural mid-western county in southeast
Nebraska” (Nevitt and Lundak 2005). That might not be
the best model for the behaviour of my study population
of Czech ornithologists who typically enjoy their time not
being ‘arrested’. Second, self-reporting (Dudley 2002) is a
standard methodology in studies of alcohol consumption.
Sheil et al. themselves cite a paper that claims “self-reported
alcohol intake generally appears valid” (Mukamal et al.
2008). Third, I know most survey participants personally.
If a teetotaller colleague of mine claimed he drank dozens
or hundreds of beers last year I would know he was lying. If
a colleague of mine who spent a dozen of evenings in a pub
with me last year claimed he drank just 50 litres per year, I
would know that he was lying too. Intimate knowledge of
the study organism and population is crucial.

The original paper contains data suggesting that re-
spondents did not pull their answers out of thin air:
“changes [between 2002 and 2006] in per-capita con-
sumption were very small”. ‘Within-respondent reliability’
estimated by Pearson correlation (Martin and Bateson
2008, p. 74–78) is 0.95. Could students of molluscs or
ducks dream about a better consistency in behaviour of their study subjects?

Papers that criticised the self-reported beer study were also based on studies that were self-reported (Moya-Laraño 2008, Sheil et al. 2008). But the authors did not acknowledge that. The effect on the reader is predictable – the criticised study is methodologically unreliable (bad impression) and the opponents offer plausible alternative opinions flattering readers (good impression). The present critique does not refute the existence of plausible alternatives. But it does undermine the robustness of the criticism of ‘self-reporting’ methodology because 1) there are no non-self-reported alcohol studies that might serve as a baseline for judging the reliability of self-reported studies (Dudley 2002), and 2) counter-arguments by opponents of the beer study rest on the very same methods as the beer study – by casting doubts on the beer study they cast doubts on themselves. You cannot have it both ways. Either self-reporting is not a good method – and then both the beer study and all ‘moderate drinking is good for health’ papers (add also all public opinion polls!) go to a litter bin. Or self-reporting is reasonably reliable and then both former and latter studies are acceptable.

**Direction vs shape of the relationship**

Sheil et al. note that “there is no theory implying a strictly linear relationship.” I agree. I also stress that this is not an argument against the beer study. The paper is neither about the shape of the relationship, nor does it claim anything about “a strictly linear relationship” as wording by Sheil et al. implies. I predicted a “negative correlation”. The potential finding of a non-linear shape of the consistent unidirectional relationship would not detract anything from the conclusion of “negative correlation” – it would only add extra details. Similarly, fitting a logistic growth curve to growth data does not change the conclusion that there is a positive correlation between an individual’s age and mass. In parallel, there is no discrepancy between Grim (2008) and Sheil et al. (2008). Sheil et al. (2008) illustrate a common logical error – the straw man (Pirie 2006, p. 155–157).

Moya-Laraño (2008) substantiated his arguments using statistical analysis. He tested whether there was a difference between the slopes of the beer vs publication relationships for “moderate” and “heavy” drinkers. He applauds his rigorous approach but there are two – understandable – confusions in his re-analysis.

The minor point is that Fig. 1 does not show “two data sets” as all data are for year 2006. Full and open circles differentiate between “past” researchers (that had been included in both 2002 and 2006 censuses) and “present” researchers (that started to publish only after my first survey and were therefore included only in 2006). I hoped this would be clear from the methods-section, but it was not. I apologise for this confusion.

Moya-Laraño assumed that the breakpoint between moderate and heavy drinkers lies in the middle of the data set (here I assume that Moya-Laraño accepts the medical definition of ‘moderate’ and ‘heavy’ drinking, i.e. the definition accepted in all papers on the issue he cited). I intentionally used Box-Cox transformation, thus it was not possible to determine from the presented data where the borderline between moderate and heavy drinkers was. However, one could expect that ‘heavy’ drinkers would not be frequently included in the sample of professional scientists. Indeed, only 4 (out of 34) persons in the data set could be classified as ‘heavy drinkers’ (standard definition of ‘moderate drinking’ = up to 2 standard drinks per day, ‘standard drink’ = 12 g of alcohol, a beer with 4.5% of alcohol, Dufour 1999). Recalculating the data from Fig. 1 excluding the heavy drinkers shows a significant decrease in publication performance in low and moderate drinkers (b = −0.31, SE = 0.12, p = 0.018). There is no evidence that could endorse “a break to moderate drinkers”.

**Positive effects of alcohol on creativity: I drink, therefore I die – alcohol and n(d)umbness**

“[Scientific] ideas are blossoms of virtue that fail to open their petals and wilt quickly in the fumes of boisterous partying.” Ramón y Cajal [1897] (1999, p. 101)

According to Sheil et al. alcohol is “boosting creativity”. In the first paragraph they claim that “alcohol opens minds and promotes originality” and “this idea gains credibility from the creative arts … generally”. Set aside the fact that they refer to correlative data which, in their own words, “prove[s] little”. Set aside the fact that creativity in arts does not equal creativity in science – one term can have more than one meaning. Or the fact that they refer to papers based on anecdotal and self-reported evidence. Or even the fact that they contradict themselves by first stating that “It is well established that excessive alcohol consumption is bad for mental and physical functioning” and then, when discussing benefits of “moderate drinking”, they cite papers on problems of alcoholism in artists.

They refer to two papers — Beveridge and Yorston (1999) and Tolson and Cuijet (2007). Both papers discuss a view that the beneficial drinking idea “gains credibility from the creative arts”, but both conclude that the credibility is misplaced. Sheil et al. draw readers’ attention to the former, but remain silent about the latter.

In the second paragraph, Sheil et al. turn to effects of alcohol on scientific “inspiration”. They report “one story” about beer being used in an experimental apparatus and not for drinking. When discussing general patterns, it is logically fallacious to use anecdotal evidence – the fallacy of secundum quid (Pirie 2006, p. 145–147). Such ad hoc information only optically ‘boosts’ the size of the section on creativity, so the incautious reader might be mislead into
believing that the arguments supporting the claims of Sheil et al. are better than what, in reality, they actually are.

In the third paragraph, Sheil et al. cited some studies that found alcohol “to facilitate some creative functions and inhibit others, with ‘significantly higher levels of originality’ being the overall outcome in many cases”. However, science is more than ‘originality’; scientists spend the most time with routine work, i.e. collecting data using established methods, analysing data using routine statistical procedures, or writing papers in “the predictable, stilted structure and language” (Sand-Jensen 2007). The studies that Sheil et al. cite were short-term experiments providing no evidence for long-term effects – and scientific work is not a short-term thing. Finally, Sheil et al. wrote that the “overall outcome” was “significantly higher levels of originality”. But their source states the opposite: “there was, however, no difference between groups regarding scientific value” of their work (Norlander 1999). As the author of papers that Sheil et al. cited notes: “Alcohol both enhances and disturbs creativity. Truly enough, the creative process includes very different elements. Creativity is not only about the grand ideas.” (T. Norlander, pers. comm.).

Originality is of little value when other necessary aspects of a scientist’s work are deficient. Scientists with lots of original ideas might be forgotten when s/he is unable to do the routine work. In contrast, a non-original scientist can perform relatively well – s/he can still employ the methods and designs of others and test their original hypotheses. Most scientific papers are done this way by ‘copying’ (O’Connor 2000).

In summary, studies cited by Sheil et al. provide equivocal, tangential or even opposing evidence for their views. But I do not doubt that many readers were persuaded – namely those readers who knew what to think before they even started to read (Lord et al. 1979).

Are there any positive effects of alcohol on human health?

Both Moya-Laraño and Sheil et al. claim that alcohol in small amounts improves human health. First, possible benefits of low/moderate drinking do not contradict the fact of the generally negative effects of alcohol across the whole spectrum of drinking from teetotaling to heavy boozing. Unless a particular range of values of a variable is specified, we, by default, mean the whole natural range – just like I did in the beer study. The correct method is to first test the widest value range of the factor being assessed (Kamil 1988, Grim 2005). Second, almost every negative factor is beneficial when applied in small doses and every medicine kills in excessively large amounts. Third, alcohol in small doses improves only some particular health parameters, not the health in its entirety – even moderate drinkers show impairments of health and cognition (reviewed by Eckardt et al. 1998, Rehn et al. 2003).

In his discussion, Moya-Laraño cites 10 studies purportedly supporting positive effects of alcohol on health. Moya-Laraño starts with “moderate drinking” but then suddenly switches to “Mediterranean diet”.

First, Moya-Laraño does not explain why he switched from apples (effects of alcohol) to oranges (effects of diet). Second, he does not explain why he switched from other apples (publication output) to other oranges (human health). Could it be that healthier scientists produce better science? Perhaps yes, but Moya-Laraño does not give any evidence. Third, papers cited by Moya-Laraño can be summarised as follows: 1) all 10 studies are on diet and thus are irrelevant to the discussion of the beer study, 2) some papers studied effects of the Mediterranean diet on mortality (not on health), 3) most do not discuss effects of alcohol at all, 4) some show that effects of alcohol are non-significant, 5) some even discouraged drinking, 6) some of them show poor evidence for wine drinking on health (sometimes negative effects), 7) effects sometimes disappear when confounding variables are controlled (the results of analyses without confounders are then invalid), 8) subjects of the cited studies were exposed only to a combination of wine with the Mediterranean diet – this ‘interactive’ design is insufficient to say anything on the possible effects of beer drinking which is accompanied by a non-Mediterranean diet in a non-Greek population in my study area.

Moya-Laraño’s arguments succeed (partly) in proving that some special kind of diet is good for health; i.e. he argued completely besides the point of how alcohol affects cognition. This is “one of the oldest [argumentative] fallacies known to us” – ignoratio elenchi (i.e. the fallacy of irrelevant thesis; see Pirie 2006, p. 94–97). To summarise, Moya-Laraño provides no evidence that could be used as an argument in the current debate in principle.

Sheil et al. claim that “moderate regular alcohol consumption provides significant health benefits” and they cite Mukamal et al. (2008). Not only does the paper not address the issue at all – it is a sociological paper about “how physicians and patients react” (p. 188) to the hypothesis of alcohol related health benefits. Such a paper is interesting in its own right, but is tangential as an argument about the health effects of alcohol. Analogously, one could cite a poll showing that more than half of United States citizens disagree with the evolutionary theory (Miller et al. 2006) as evidence that “Darwin was wrong.”

Mukamal et al. (2008) start the abstract: “The relationship of moderate alcohol use and health remains controversial and uncertain.” They follow: “Even moderate alcohol consumption can have risk. ... increase in breast cancer risk associated with moderate alcohol use. ... the increase in cirrhosis attributable to even moderate alcohol intake... .”

Further, Sheil et al. write that “moderate alcohol consumption lowers the risk of cardiovascular problems and increases overall life expectancy”. But is the effect size big enough to over-compensate for the negative effects of alco-
The answer seems to be no (Goldberg 2003, Jackson et al. 2005 and references therein).

Sheil et al. and Moya-Laraño illustrate that “citation malpractice is widespread” in ecology; unfortunately, there are many instances where the support of an assertion by the cited reference proves to be ambiguous, non-existent, or even contradictory (often we only notice this when our own work has been mis-cited! ... whatever the reasons, this does not change the fact that the final reader of the article will be misled” (Todd et al. 2007).

What is the general lesson for us? Never cite ‘blindly’, never judge papers according to their abstracts only. An abstract cannot reveal that a study is wrong – the abstract does not show misapplied statistics, insufficient sample sizes or erroneous assumptions (Yudkin 2006). Always do check original sources – no matter whether they are cited by your opponents or yourself.

Selective evidence

Considering the benefits of any activity in any organism is meaningless unless the costs are considered too. A major flaw of criticisms of the beer study is listing the benefits of moderate drinking but ignoring the costs and their final pay-off.

As Jackson et al. (2005) concluded “Any coronary protection from light to moderate drinking will be very small and unlikely to outweigh the harms. While moderate to heavy drinking is probably coronary-protective, any benefit will be overwhelmed by the known harms. ... Do not assume there is a window in which the health benefits of alcohol are greater than the harms.” See also Goldberg (2003).

Although I agree that authors can explicitly state that they are going to present evidence for just one point of view, to counterbalance the other one (Sheil et al. 2008), they should be clear about what conclusions do the majority of evidence support. Otherwise they commit the fallacy of one-sided assessment (Pirie 2006, p. 121–123).

The ‘correlation is not causation’ mantra

“Third causal factors could explain the observed correlation, i.e. underlying factors of social and economic origin could explain both heavy drinking and less success in high-level research and publishing – in principle, without any causal link between the two outcomes” (Sheil et al. 2008). The first part is true. Sheil et al. just overlooked and/or forgot to stress that it also holds for virtually all studies of 1) positive effects of drinking on human health, 2) studies of publication bias, and 3) studies of creativity related to alcohol consumption that were cited by them and by other opponents of the beer study. The second part is ambiguous. Medical and psychological studies, both relative and experimental, showed overall negative effects of drinking on human performance. Therefore, there is little doubt whether there is a negative causal link between drinking and scientific performance – it inevitably follows from the physiological and cognitive traits shared by all individuals of the Homo sapiens species. The only question is whether the relationship is detectable without experimental manipulation.

Importantly, “The data on alcohol and cardiovascular disease are still correlative, whereas the toxic effects of alcohol are well established” (Goldberg 2003). Dudley (2002) observed that there was not a single study that would support the “positive effects of alcohol” experimentally. “Several concerns have been raised about the studies that link moderate alcohol use and [coronary heart disease]. Most importantly, all of these studies have been observational, raising the possibility that an unmeasured confounding factor explains the association” (Mukamal et al. 2008).

In general, the use of causative language (‘affects’, ‘causes’, ‘influences’) is ubiquitous in ecological studies based on observational/correlative data! Studies of climate change-driven ecological changes (Parmesan 2006), biodiversity gradients (Fuhrman et al. 2008) and ecological communities (Novák and Konvička 2006) are just examples of many sub-fields in ecology that unscrupulously use causative language, or loosely switch between the language of correlation and that of causation. I do not question the validity of those studies. I just stress that, in this respect, the beer study was more cautious than a typical ecological study.

I wonder what kind of long-term experiments the critics of correlational approach would suggest to do. It is unethical to randomly assign say a couple of hundreds of scientists into treatments ‘teetotallers’, ‘moderate drinkers’ and ‘heavy boozers’ and thus ruin the careers and live(r)s of the latter. Not surprisingly, humans are studied without long-term experimental manipulations by a “primarily correlational approach” (Danchin et al. 2008, p. 690).

Alternative explanations: drowning of the sorrows

Every finding is open to more explanations (Sheil et al. 2008). Existence of other explanations per se has little to do with the validity of any hypothesis. I do not refer to the fact that until those non-compatible alternatives are supported by better evidence than the current hypothesis, they cannot reject it. They can merely cast doubts because for any phenomenon there are multiple hypotheses. Also, current absence of suggested alternatives does not mean we have reached the final solution – nobody knows what new ideas will see the light of day tomorrow.

Let’s consider an alternative explanation “researchers drown their sorrows after their papers are rejected” (Sheil et al. 2008). This reverse causation hypothesis is, in prin-
biases and prejudices in both scientists and laymen. This use the fallacious arguments etc. The differences between reactions to the study illustrate common fallacies committed by both scientists and laymen: using selected evidence, mis-citing source materials, using logically fallacious arguments etc. The same correlative approach is unreliable (or outright wrong) when it leads to conclusions we do not like. Sadly, results that are because of reasons un-related to science expected or wanted are blindly accepted. In contrast, results that are unpleasant are either ridiculed or rejected outright without any counter-evidence (Lord et al. 1979).

We humans seem to detect drawbacks only when we can criticise them elsewhere, e.g. in the work of other scientists (Sheil et al. 2008). This provides a clear Jungian impetus for our own work – let us pay more attention to the possibility that what we criticize in the work of others essentially reflects faults in our own work. In doing so, we can improve the quality of ecological research – and our everyday communication – in the future.

Acknowledgements – I am grateful to L. Wei (Hainan Normal Univ.), J. Rutila (Univ. of Joensuu) and M. E. Hauber (Univ. of Auckland) for their hospitality. Many thanks to T. Albrecht, P. Cassey, A. Dvorská, P. Jansen, P. Kovafik, M. Krist, C. Lottie, A. P. Møller, B. Pinshow, V. Remeš, D. Sheil, D. Storch, K. Weidinger and others for discussions of this essay. D. Campbell corrected the language. My research is supported by Human Frontier Science Program Award.

References


Conclusions – the beer study as a mirror for scientific community

“Why do you see the speck in your brother’s eye but fail to notice the beam in your own eye?” (Matthew 7: 3)

In sum, the beer study is an ordinary correlative study – just like many ecological studies and virtually all publication bias studies. It showed that the mere change of a studied subject (cf. methods) can lead to a strong criticism of standard research practices in a large and established discipline.

The criticisms of the beer study illustrate common fallacies committed by both scientists and laymen: using selective evidence, mis-citing source materials, using logically fallacious arguments etc. The differences between reactions to the beer study and ‘normal’ ecological studies – that use the same methods, sample sizes, statistics etc. – expose biases and prejudices in both scientists and laymen. This highlights the general problem of double standards. For example, a correlative approach (self-reported data, sample size of a couple of dozen data points – fill in what you like) is fine when results support our prejudices. But the very same correlative approach is unreliable (or outright wrong) when it leads to conclusions we do not like. Sadly, results that are because of reasons un-related to science expected or wanted are blindly accepted. In contrast, results that are unpleasant are either ridiculed or rejected outright without any counter-evidence (Lord et al. 1979).

We humans seem to detect drawbacks only when we can criticise them elsewhere, e.g. in the work of other scientists (Sheil et al. 2008). This provides a clear Jungian impetus for our own work – let us pay more attention to the possibility that what we criticize in the work of others essentially reflects faults in our own work. In doing so, we can improve the quality of ecological research – and our everyday communication – in the future.

Acknowledgements – I am grateful to L. Wei (Hainan Normal Univ.), J. Rutila (Univ. of Joensuu) and M. E. Hauber (Univ. of Auckland) for their hospitality. Many thanks to T. Albrecht, P. Cassey, A. Dvorská, P. Jansen, P. Kovafik, M. Krist, C. Lottie, A. P. Møller, B. Pinshow, V. Remeš, D. Sheil, D. Storch, K. Weidinger and others for discussions of this essay. D. Campbell corrected the language. My research is supported by Human Frontier Science Program Award.

References


Pirie, M. 2006. The use and abuse of logic. – Continuum.