

Reviewer's report

Title: Neuropsychological effects of chronic low-dose exposure to polychlorinated biphenyls (PCBs)

Version: 1 **Date:** 21 April 2005

Reviewer: David Bellinger

Reviewer's report:

General

Given that this study is described as exploratory (since it is inappropriate not to detect subtle differences at this stage of research), the Bonferoni correction for alpha renders this more of a confirmatory study, increasing the likelihood of overlooking subtle associations. This risk is increased, as well, by the very modest sample size. Furthermore, some of the effect sizes for the associations that did not meet the corrected alpha level were substantial, considerably larger than those typically observed in studies of low-dose chemical exposures. As indicated in the paragraph Dose-response relationships in the Results, some of the correlations observed between PCB28 and behavioral variables were as high as 0.54, with several others >0.3 . These are extremely high correlations for multi-determined, complex outcomes such as performance on a neuropsychological test. The effect sizes listed in Tables 3 and 4 frequently exceed 0.7 and range as high as 1.38 for d_2 . A value of $d < -0.2$ was used as the criterion for low performance on a behavioral variable, and inspection of the d s in Table 4 shows that values for all but a few variables were this extreme. (Although I did wonder how a difference of 0.2 in the group T score means on WAIS Similarities, for which the SD was 9 in both groups, could result in a d_2 of 1.38.) The alpha correction used leads the authors to minimize the potential importance of these correlations and effect sizes, as does their use of the criterion of clinical impairment at the individual level to judge the importance of the associations (statements such as, [differences] which, however, were within the range of the normative population and differences between exposed subjects and non-exposed controls were relatively small and because mean normative values did not exceed the levels required for the identification of impaired persons. Applying an individual criterion for impairment to the interpretation of group differences on a population basis can be highly misleading, as I argued in a recent paper in Environmental Research (2004) entitled, What is an adverse effect.

Specific Comments

1. What considerations resulted in the decision to have a sample size of 60? Were any selection criteria (inclusion/exclusion) applied? Was the number limited by the number of teachers working at the exposed school? Reference is made at the end of the Discussion that the associations, might have been moderated by potential confounders such as differences between city and rural populations in life style and nutritional habits The Methods indicates only that teachers were drawn from two high schools near Heidelberg. More information needs to be provided about these potential differences. Also, the statement that the group differences might have been due to different performance profiles associated with school type, or motivational factors needs to be explained more. Why would performance profile differ by school (in the absence of an effect of PCB exposure), and what motivational factors might have varied between schools (and why)?
2. Were the neuropsychological evaluations conducted in a blinded fashion or did the tester know the school at which an examinee worked?
3. The MS indicates that no teacher had a history of neurological or psychiatric disorder. It seems

unlikely that, in a sample of 60 adults, none had a history of, in particular, a psychiatric disorder. How was this information collected, and what criteria were applied?

4. Additional information about the measurement of the concentrations of PCB in air samples would be helpful. How were the samples collected? Were they 24-hour samples? When were they collected in relation to the participants blood PCB levels? Near the end of the section, External Exposure, please clarify the statement that An accumulation of potential risk factors could not be confirmed for either group.

5. What was the temporal relationship between blood sampling and potential PCB exposure in the school? Given that exposure appeared to consist primarily of low chlorinated congeners with shorter half-lives, some exposure misclassification could have occurred if exposure and blood sampling were separated by a few months. Also, I am not clear on the implications of the fact that the GC-ECD was replaced by GC-MS. Why not just say that the samples were measured by GC-MS, if those are the levels used in the analyses?

6. The MS indicates that a total toxicity index was not calculated, since complete information about all congeners was not available. No study ever measures all PCB congeners, yet TEQs are estimated, so this cannot be a sufficient reason not to attempt to do so here.

7. What is the basis for the statement that computerized testing was added to increase the sensitivity of the battery in the area of executive and attentional functions? Why is this statement made specifically with reference to executive and attentional functions as opposed to other neuropsychological functions? What evidence can be marshaled to indicate that computerized testing is, in fact, more sensitive than examiner-administered tests for these domains?

8. How the preceding factor analyses of control subjects data were conducted and how it structured the analyses is not clear. These analyses are not mentioned in the Results section.

9. Why gender received so much attention as a potential confounder is puzzling, especially since it was unrelated to PCB exposure variables and behavioral outcomes. Given the small size, the degree of freedom used up by adjustment for gender might have been used more productively adjusting for some other factor that actually functioned as a confounder in this data set. This section does indicate that, in separate analyses, adjustments were made for openness and alcohol consumption on self-report and estimated intelligence level and alcohol on behavioral measures. Why these variables and why only for these outcomes? How were they identified as potential confounders. Table 1 shows that the two groups did not differ in IQ or alcohol (although the latter variable was clearly highly skewed).

10. Although the Data analysis section indicates that a corrected alpha of 0.004 was applied, the Results refers to significant associations for which P-values were >0.004 . Also, given the study design (matching exposed and controls on various factors), were the analyses conducted using methods appropriate for matched data?

11. Greenwald et al. is not included in the References section.

12. It isn't clear what is being shown in Figure 2. Are these values only for PCB 138, 153, and 180, as the legend suggests? If so, this doesn't seem very helpful insofar as the school-related exposure was determined to consist mostly of congeners 28 and 101. If PCB28 levels are reflected in Figure 2, how were the large number ($>90\%$) of controls with PCB28 below the detection threshold represented?

13. In the first paragraph of the section, Neuropsychological results, I think the reference to Table 1 should be Table 3? Similarly in the next paragraph in this section, should reference to Table 3

instead be Table 4?

14. In Tables 3 and 4, a delta is used to indicate a potentially relevant effect and refers the reader to the text for an explanation, but I could not find such an explanation.

15. The self-report measures that differed between groups defined on the basis of external exposure (well-being, distractibility of mental processes, introversion) seemed to differ from those that were associated with biological measures of PCB (tiredness and slowing, emotional reactions). Can the authors comment on potential explanations for this? Does it suggest that the associations might be the result of chance?

16. The last sentence of Results states that differences were found on TAP phasic alertness and response shifting, yet the P-values are listed only as n.s. Why is attention drawn to these non-significant results and not to others?

17. The paragraph in the Discussion on the generalizations permitted (last paragraph in the section entitled External and internal PCB exposure) is quite confusing. First, it says that generalization does not seem warranted. Then, At least partial generalizations to other PCB exposure conditions might be possible. Then the last sentence of this paragraph suggests that generalization might not be possible because, A diverging PCB pattern or additional pollutants at another site may thus produce a different behavioral effect.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

I would consider that all of the comments and suggestions above need to be addressed.

Discretionary Revisions (which the author can choose to ignore)

What next?: Accept after minor essential revisions

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests.