

Letters to the Editor

re: WARD, S., 1999, An investigation into the effectiveness of an early intervention method for delayed language development in young children. *International Journal of Language & Communication Disorders*, 34, 243-264.

Dear Editor,

Phenomenal results or practice effect?

Ward presents very provocative results because (1) the randomized group design is a strong one, (2) she includes a long-term follow-up of the effects, (3) the families are from low socio-economic levels and (4) the size of the treatment effects are larger than any I have seen in the communication disorders literature. This letter is devoted to understanding the latter attribute.

The sizes of the treatment effects were not computed in the paper and the information given is ambiguous. Therefore, the actual size of the effects are unclear. It is not clear whether the appropriate standard deviation (SD) needed to compute the effect size is the SD of the study's sample (as in the case of language age/chronological age quotients) or the SD determined by the transformation process (as in the case of standard scores for standardized tests, usually 15). The text of the paper indicates that 'language quotients' (which are typically between 0 and 1.0) were used, but reports whole numbers. If the more liberal SD is used, the effect sizes range from 2.0 to 3.0. If the more conservative SD is used (i.e. SD = 15), the effect sizes range from 1.87 to 2.1.

Both metrics (i.e. language quotients and standard scores) are indices of the degree of language delay with respect to the chronological age (e.g. IQ), not the developmental level (e.g. mental age). It is particularly difficult to show treatment effects on indices of degree of delay because one must show greater increase on, not just skill level, but rate of development. I have not seen even the lowest of the more conservative estimates of the effect size (i.e. 1.87) on indices of language delay in any other well-conducted treatment study in the communication disorders literature.

Two possible explanations for the unexpected effects can be ruled out. First, we ruled out the possibility of intentional, dishonest reporting on principle. Honest reporting of one's results is the cornerstone of our field's ability to disseminate scientifically defensible and useful information. Second, the intervention methods, the measures and the intensity of treatment were not unusually strong. Ward used an intervention approach based on the principles derived from middle-class Western mothers of young children whose language skills develop unusually fast. This approach has been used in several past studies (e.g. Girolametto *et al.* 1996, Wilcox and Shannon 1998). In fact, both the Wilcox and Girolametto treatments are a more intensive application of the training model and neither showed treatment effect approaching those reported by Ward. The outcome measure Ward used was

a common standardized test of language development (i.e. The Reynell Language Development Scales). Standardized tests are typically less sensitive than target-specific measures to treatment effects because of their global nature. The Reynell has been used as an outcome in other treatment studies with less impressive results (e.g. Yoder and Warren 1999).

A third possible explanation is that intervening with such young children (mean age of 10.6 months at the start of the intervention) results in unexpectedly large effects. Because the children in the Ward study were younger than virtually any sample in treatment studies in the communication disorders literature, I consulted the intervention literature for at-risk infants (< 12 months) to examine this hypothesis. The relevant effect sizes were those on measures of degree of delay. After reviewing the literature on interventions with at-risk infants, Blair and Ramey (1997) concluded that the treatment group had indices of developmental delay that were between 0.36 and 0.61 of an SD greater than the control group. However, the authors of a meta-analysis of the intervention literature for children birth to 3 years concluded that the earlier the onset of treatment the better the outcome for children with mild disabilities (Shonkoff and Hauser-Cram 1987). One group treating high-risk infants in and just out of the neonatal intensive care unit showed an effect size of 1.6 (Als 1997). It should be noted, however, that both the studies Blair and Ramey reviewed and that of Als had treatments that were quite a bit more intense than that in Ward. On balance, it seems reasonable to conclude that it is possible that the early age of the Ward participants is one explanation for her large effects.

However, there were some design problems that may explain why the magnitude of effects reported in Ward were larger than even more intensive treatments with even younger infants (e.g. Als 1997). Not all design problems are addressed adequately by random assignment. For example, testing effects, differential attrition and examiner bias can occur even when children are randomly assigned to treatment groups.

In the Ward study, only the treatment group was given language tests monthly (i.e. every intervention session). This was done because the treatment was to end when the children were within normal limits for their age. It was not clear whether the Reynell was the instrument given. But even if the REEL was given instead, children tend to become better at taking tests with practice. This has been found to be true even when feedback on performance is not given (Huck and Sandler 1979).

It is stated that attrition was 17 and 16% each year, but it is not stated whether more attrition occurred in the control group than in the treatment group. If so, and if the participants who dropped out of the study were those least needing treatment, then the magnitude of the effect may be inflated.

Finally, the authors state that the examination was not blind to treatment assignment. This can bias examiners unconsciously to alter the administration procedures in a way that favours the experimental group (Huck and Sandler 1979). However, it should be noted that many treatment studies suffer from the same threat to internal validity and still have not shown such large effect sizes.

However, if future research on these children shows socially important outcomes (e.g. enrollment in special education), then it is unlikely that the design problems explain the effects. There is some hint in the paper of such outcomes in the form of more control group children being referred to speech-language services than experimental group children. Even if such educationally important outcomes are

documented, an important question remains: 'do some parents have inadequate resources, sufficiently different values about parent-child interaction, and/or sufficiently delayed or different children, that they are not able to competently and consistently implement Ward's treatment over time?' If so, then the most effective treatments are likely to be treatments that match children's abilities and that address families' psychological and social needs. Finally, the unexpectedly large effect sizes in this study cry out for replication.

References

- ALS, H., 1997, Earliest intervention for preterm infants in the newborn intensive care unit. In M. Guralnick (ed.), *The Effectiveness of Early Intervention* (Baltimore: Brookes), pp. 47-76.
- BLAIR, C. and RAMEY, C., 1997, Early intervention for low-birth-weight infants and the path to second-generation research. In M. Guralnick (ed.), *The Effectiveness of Early Intervention* (Baltimore: Brookes), pp. 77-98.
- GIROLAMETTO, L., PEARCE, P. S. and WEITZMAN, E., 1996, The effects of focused stimulation for toddlers with expressive vocabulary delays. *Journal of Speech and Hearing Research*, **39**, 1274-1283.
- HUCK, S. and SANDLER, H., 1979, *Rival Hypotheses: Alternative Interpretations of Data Based Conclusions* (New York: Harper & Row).
- SHONKOFF, J. and HAUSER-CRAM, P., 1987, Early intervention for disabled infants and their families: a quantitative analysis. *Pediatrics*, 650-658.
- WILCOX, M. J. and SHANNON, M. S., 1998, Facilitating the transition from prelinguistic to linguistic communication. In A. Wetherby, S. Warren and J. Reichle (eds), *Transitions in Prelinguistic Communication* (Baltimore: Brookes), pp. 385-416.
- YODER, P. J. and WARREN, S., 1999, How intentional communication is related to later productive language in children with developmental delay? (submitted).

PAUL YODER
 Department of Special Education,
 Box 328 Peabody College, Vanderbilt University,
 Nashville, TN 37203, USA

Dear Editor,

We have read Ward's paper on her intervention programme in which she makes strong claims about how language disorders in young children can be prevented. This research will obviously be of great interest to all who wish to provide an effective and cost-efficient speech and language therapy service to pre-school children and, therefore, deserves serious consideration. However, having studied the paper carefully, we conclude that further research is required before practitioners think about changing their style of intervention in response to this study.

Our concerns fall under two main headings: the underlying theoretical motivation for the study and the methodology. We will start with our first concern. Ward makes an assumption that parents (especially mothers) are responsible for language acquisition in their children and it therefore follows that language disorders are the result of defective input or an inappropriately noisy environment. There are passing references to some very early work by Chomsky, but there is no evidence in this paper that Ward is aware of a considerable body of research based on the notion of the innateness of the child's grammar (the child's capacity to develop language). Similarly, it is not obvious that she is familiar with the advances made in specifying the diverse nature of child language disorders (the terms delay and disorder are used interchangeably).

The identification of children for the study was done on the basis of the child's failure on a screening test at 8 months. Inclusion in the study was confirmed by assessment on the Bzoch-League Receptive-Expressive Emergent Language Scale (REEL: Bzoch and League 1991). Children with a language quotient < 90 on the REEL were considered as language delayed (although parents were not informed of this), and allocated to either control or experimental groups. The implication is that these children are at risk for persisting language impairment. However the research literature has always suggested that there is wide range of normal variation in language development at this age. A large scale study by Bates *et al.* (1995) of 1803 children between 8 and 30 months of age shows this variation to be very wide indeed (the REEL, as Ward says, is based on a sample of only 50 children). To take receptive vocabulary as an example from the Bates *et al.* study, at 10 months the 1.28 standard deviation range (accounting for 80% of the data) goes from zero words understood to 144. The authors note that their sample comes from a wide socio-economic range, but is heavily weighted towards middle class families.

It may be the case that a proportion of the children included in Ward's study would have gone on to have a normal profile of language skills at a later stage. Law *et al.* (1998) report that 'spontaneous remission of speech and language delays is high in the pre-school period' (p. 16), although it must be said that this applies only to expressive delays. Do Ward's procedures identify children who have language delay, do they identify children who have no current delay but are at risk, or is she just identifying a group of children within the range of normal variation for this age group? The answers to these questions have important implications for service delivery.

A second concern is to do with the nature of the intervention. The advice given to parents and demonstrated during home visits is based soundly on studies of normal adult-child interaction. The problem though is that, because studies of such interaction looked at what a majority of adults (in Europe or the USA at any rate) do naturally, we are left with the puzzle of why these particular adults need this advice. While there is evidence that adults may vary in the effectiveness of their interaction with children (Harris 1992, for example, links making comments contingent on the child's focus of attention with her faster developing group), there are no studies that we know of that show a link between quality of interaction and language delay/disorder. In fact, in the current study, Ward gives no evidence that the children's linguistic environment deviates from the norm. It appears to be an assumption made because the children are delayed in language development. There is some suggestion that socio-economic background may be a factor here, but if this is the case then comparison intervention studies need to be carried out with families from different socio-economic backgrounds. The existence of families in which there are both language-impaired and non-language-impaired individuals is good reason to be cautious about apportioning too much importance to environment. A good example is the K family as reported by Gopnik and Crago (1991).

What then are the implications here for cost-effective intervention? It would certainly be helpful for parents to be given advice on talking and interacting with their children if there is evidence that provision of this normal environment is lacking. Indeed, this may account for the encouraging results that Ward obtained when comparing her experimental and control groups. We would suggest, however, that numbers of instances where this is appropriate will be small given the basis of

studies of parental interaction which are based on what the majority of parents do. Where such advice is required, it may be more appropriate for it to come from a health visitor or other professional routinely involved with the child. The norm within many services is that despite the wide range of normal variation and consequent difficulties of diagnosis, many children are referred to speech and language therapy from quite an early age (~ 2 years). Monitoring through regular review, plus giving advice where necessary, is arguably less draining on time than the intervention programme proposed by Ward. Her screening instrument, with its focus on early auditory perceptual skills, may lead to even earlier referrals of this type and subsequent more effective monitoring.

There are children who go on to have serious developmental language disorders that affect their education, social and later working lives and we urgently need instruments to help the identification of such children as early as possible. Equally, we need studies that will give some indication as to when remediation should be targeted to be most effective. Unfortunately, this paper addresses neither of these needs.

References

- BATES, E., DALE, P. S. and THAL, D., 1995, Individual differences and their implications for theories of language development. In P. Fletcher and B. MacWhinney (eds), *The Handbook of Child Language* (Oxford: Blackwells), pp. 96–151.
- BZOCH, K. and LEAGUE, R., 1991, *The Bzoch-League Receptive-Expressive Emergent Language Scale*, 2nd edn (Texas: PRO-ED).
- GOPNIK, M. and CRAGO, M. B., 1991, Familial aggregation of a developmental language disorder. *Cognition*, **39**, 1–50.
- HARRIS, M., 1992, *Language Experience and Early Language Development: From Input to Uptake* (London: Lawrence Erlbaum).
- LAW, J., HARRIS, F., BOYLE, J. and HARKNESS, A., 1998, *Screening for Speech and Language Delay: A Systematic Review of the Literature. Summary of Main Findings* (York: NHS Centre for Reviews and Dissemination).

CAROLYN LETTS and SUSAN EDWARDS
 Department of Linguistic Science, University of Reading,
 Reading RG6 6AA, UK

Dear Editor,

The findings of Dr Sally Ward, if replicated, must rank among the most cost-effective early stimulation programmes yet devised. Against this background, all possible alternative explanations for her findings must be explored and the implications need careful consideration. First, health visitors must have been very skilled at identifying infants whose only developmental problem was in their linguistic skills. When these children were followed up by therapists at home, most were within normal limits on other developmental measures and only showed problems in their language skills. This accuracy is remarkable considering that the health visitors used only a simple language screening tool and did not carry out a rigorous developmental examination.

Second, the language screening process was unusually efficient. The paper gives a figure of 6% false-positives, but the number examined at home after the initial

screening was 182, of whom 37 were deemed to be false-positives. This is $\sim 20\%$, but still represents a remarkably good screening process.

Even more striking is the apparent predictive power of this process. If we concentrate only on the control group who it is assumed received no significant intervention, it appears that this simple screening for linguistic delays in the presence of otherwise normal development effectively identified a group of children with a similar profile some 2 years later. This indicates a degree of stability in developmental trajectory, which is greater than most people would have expected. However, there is some support for Ward from another study on the continuity of communicative competence (Smith 1998).

Turning now to the intervention, it is not implausible that such a simple intervention should be so effective, but the improvement is more dramatic than one might have expected from the literature on other early intervention programmes.

Before accepting these results, it is important to look at any possible design errors or bias. The first and most obvious is described by Ward—the children were assessed by the intervention team who could not be blinded to the status of each child. I doubt that the results could be explained solely on the basis of observer bias, but perhaps the children performed substantially better with the therapists in the experimental group simply because she was a familiar figure, whereas for the control group the children were being seen in an unfamiliar situation by someone they did not know. This does not seem a likely explanation but only the author can say if it is plausible.

The second point is the hypothesis under test. At one point in the paper the outcome being considered is a fall in the number of children referred to the speech therapy service. How did these referrals originate? There is room for bias in the extent to which children were referred by the therapists in the research team. Furthermore, if the regular health visitor knew that a child was in the treatment group, she might be less likely to make such a referral.

The third difficulty in interpreting these results is in understanding the numbers of subjects. We need a diagram illustrating the total number of children screened by the health visitor, and the stages at which children were rejected from the study or dropped out. The health visitors must have screened out 20%. This is a substantial number of children and represents the low normal end of the distribution, rather than children with 'specific' language impairment, where only a few per thousand are affected. This is important because of recent evidence that environmental influences may play a large part in the variation between children in language acquisition within most of the range but genetic effects may be more important in the bottom few per cent (Dale *et al.* 1998). Thus, if Ward's results are replicated, they might raise the overall level of language skill but not reduce the number of children with more severe problems.

It is difficult to work out exactly how the randomization process occurred. It would be very easy to include only parents who expressed an interest in the study. One would expect parents to be fully informed about the study, to give signed consent to take part and for the name then to be randomly allocated to the treatment or control group. The slightest opportunity for the researchers subconsciously to influence this process could result in a more compliant or interested group of parents being in the experimental group and this could explain the eventual findings.

Tables 6–8 do not indicate the number of children involved. Those children who were not responding well to the intervention programme might have gradually

dropped out of the treatment group, so that the attrition process in these children would result in a bias in favour of those who had done well with the intervention (or had improved spontaneously anyway) thus boosting the results in this group—whereas the control group would have continued to show attrition on a purely random basis.

What should health authorities and provider units do now? It is unlikely that the intervention programme put forward by Ward could do harm and it may well be beneficial, but there is an obvious need for replication of the study. However, several questions should be considered before extending investment in this approach. First, does this simple intervention need trained therapists? Second, why limit this approach to the bottom quintile—might it help all children? Third, it might be important to link this with pre-reading programmes for maximum benefit (Whitehurst 1997). Fourth, what exactly are the aims? Ward's approach might help most children to accelerate language development, but will this gain be maintained and will it result in better progress in school and in particular with reading? If, on the other hand, the aim is the prevention of severe specific speech and language impairment, it is unlikely that Ward's approach will prevent such problems occurring—but because they are rare, a huge and expensive study would be needed to answer this question.

References

- DALE, P. S., SIMONOFF, E., BISHOP, D. V. M. *et al.*, 1998, Genetic influence on language delay in two-year-old children. *Nature Neuroscience*, **1**, 324–328.
- SMITH, L., 1998, Predicting communicative competence at 2 and 3 years from pragmatic skills at 10 months. *International Journal of Language and Communication Disorders*, **33**, 127–148.
- WHITEHURST, G. J., 1997, Language processes in context - language learning in children reared in poverty. In L. B. Adamson and M. A. Ronski (eds), *Communication and Language Acquisition—Discoveries From Atypical Development* (London: Jessica Kingsley), pp. 233–265.

DAVID M. B. HALL
Division of Child Health, University of Sheffield,
Sheffield Children's Hospital,
Sheffield S10 2TH, UK