Rhyme and Reading: A Critical Review of the Research Methodology

Article *in* Journal of Research in Reading · February 2002

DOI: 10.1111/1467-9817.00156

1 author:

Bonnie Macmillan
www.readin1week.com
1 PUBLICATION **48** CITATIONS

SEE PROFILE

Journal of Research in Reading, ISSN 0141-0423
Volume 25, Issue 1, 2002, pp 4–42

Rhyme and reading: a critical review of the research methodology

Bonnie M. Macmillan

School of Psychology, University of Birmingham, UK

There is debate over whether children's early rhyme awareness has important implications for beginning reading instruction. The apparent finding that pre-readers are able to perform rhyme tasks much more readily than phoneme tasks has led some to propose that teaching children to read by drawing attention to rime units within words is 'a route into phonemes' (Goswami, 1999a, p. 233). Rhyme and analogy have been adopted as an integral part of the *National Literacy Strategy* (DfEE, 1998), a move which appears to have been influenced by three major research claims:1) rhyme awareness is related to reading ability, 2) rhyme awareness affects reading achievement, and 3) rhyme awareness leads to the development of phoneme awareness. A critical examination of the experimental research evidence from a methodological viewpoint, however, shows that not one of the three claims is sufficiently supported. Instructional implications are discussed.

In a recent paper in this journal, Usha Goswami argues that the balance of research evidence supports the view that rhyme is causally related to reading (Goswami, 1999a). Elsewhere, she and others have pointed out that a critical examination of research methodology is also important when evaluating the research evidence in this field (Bryant, 1998; Bryant & Goswami, 1986; Goswami & Bryant, 1990; Goswami & East, 2000).

Accordingly, the present investigation seeks to evaluate the rhyme and reading research from a methodological perspective using a similar approach to that adopted by Troia (1999), who evaluated phonological awareness training studies. However, the present evaluation spans more than one topic, and differs from Troia's work in examining not only training studies but correlational research as well.

Twenty years ago, a number of studies were instrumental in provoking interest in whether or not early sensitivity to rhyme might be related to reading skill (Baron, 1977; Bradley & Bryant, 1978, 1983; Calfee et al, 1973; Cunningham, 1976). Since then, interest has rapidly accelerated. There is now a growing body of experimental research on various strands of the rhyme and reading topic to the extent that education officials and publishers are increasingly taking note.

However, it is important that conclusions drawn from the research in this field are reliable and accurate, since they do have the potential to directly determine the quality of classroom instruction. In this country, implications drawn from this research have already affected classroom practice. Now in place, is a national

programme of early reading instruction, the *National Literacy Strategy* (DfEE, 1998), which focuses on developing rhyme and analogy skills. Thus, as far as this country in particular is concerned, an evaluation of the quality of the research evidence in this field is overdue.

Three frequently stated research claims

The three research claims discussed in this paper arise primarily from the first of three 'causal connections' proposed by Goswami and Bryant (1990) in their causal connections model of reading development. The first causal connection proposed is that between rhyme awareness and reading, a part of the model that will be discussed in some detail later. But first, how is causation, the basis of this model, established?

Schult (1999) sets out a number of criteria that are necessary in order to establish the internal validity of causal connections: 1) *association*; 2) *time order*; and 3) *non-spuriousness*. Two other, external validity criteria which are not necessary but which can considerably strengthen confidence are 4) *mechanism*, and 5) *context*.

First then, since it is necessary to establish an *association* between two variables, the primary question explored is, are rhyme awareness and reading ability related? Establishing an association can be accomplished through correlational studies where rhyme and reading ability are measured concurrently. However, a strong correlation between two variables does not furnish any information about the direction of the relationship, about whether early reading ability is associated with later awareness of rhyme, or vice versa. Thus, *time order* can be established by measuring variation in the independent variable (rhyme) at some point in time before measuring variation in the dependent variable (reading). This is accomplished through correlational studies involving longitudinal measurement, where statistical analyses can establish not only whether a predictive relationship exists, but whether it still exists after removing the variance due to other contributing factors. Results from rhyme and reading studies, where the aim has been to establish association and time order, have given rise to the following research claim.

Claim 1: rhyme awareness is not only related to, but is predictive of reading ability. It is claimed that research evidence has indeed established that rhyme and reading are associated, and that rhyme awareness measured first, is related to reading ability measured subsequently (Goswami, 1995, p. 135).

While the first two criteria, association and time order, are enough to establish that a relationship exists, they are not enough to establish that a *causal* connection exists. Thus, in pursuing the issue of causation, an additional question explored is, does rhyme awareness determine reading ability? In order to answer this sort of question, it is necessary to demonstrate *non-spuriousness*. This means it is necessary to rule out the possibility that the relationship between the two variables (in this instance, rhyme and reading), if found, is not due to variation in a third variable (for example, variation in intelligence, age, phoneme awareness or alphabet knowledge). In this case, unlike correlational studies, which are non-experimental methods of data collection, a true experiment involving three key features (random assignment, manipulation of independent subject variables, and control of non-subject extraneous variables) is necessary. Results from studies attempting to satisfy this third criterion have given rise to a second claim.

Claim 2: rhyme awareness affects or determines reading ability. It has been concluded that research evidence supports the idea that rhyme awareness and reading achievement are causally connected (Goswami, 1999a, p. 219; Maclean et al, 1987, p. 255).

Although these three criteria, association, time order and non-spuriousness, are regarded as being sufficient to establish causation, confidence is considerably strengthened if two further criteria pertaining to external validity, *mechanism* and *context*, are met. First, it would help to confirm a causal connection between rhyme and reading if the mechanism by which rhyme determines reading ability can be explained. If hypotheses which propose a mechanism or a path of development (e.g. rhyme awareness leads to phoneme awareness, which in turn, determines reading ability) are experimentally tested, it is possible to further increase confidence that rhyme does indeed determine reading ability. Finally, experiments which can demonstrate under which conditions (or context) the mechanism operates can further help to eliminate uncertainty. Although space does permit a discussion of context, mechanism is an issue related to yet a third claim that has been made.

Claim 3: rhyme awareness leads to phoneme awareness. It is claimed that sensitivity to rhyme 'leads to' (Bryant et al, 1990a, pp. 429–430; Goswami & Bryant, 1990, p. 111) or makes it 'easier' to develop, an awareness of phonemes (Goswami, 1999a, p. 222).

In summary, three claims made in the research literature are:

1. sensitivity to rhyme is related to as well as predictive of reading ability;
2. there is a causal connection between rhyme and reading;
3. rhyme awareness leads to the development of phoneme awareness.

Method

Scope of this investigation

The purpose of this investigation is to try to determine the extent to which each of the three claims is supported by quantitative research evidence. What follows is a methodological evaluation of rhyme and reading studies, which find either support, or no support for each of these claims.

'Rhyme awareness' is defined here as the ability to detect the sound of rhyme in spoken words, or to generate words that rhyme with a spoken word; rhyme awareness occurs, according to this definition, only in the absence of print. In contrast, the use of the term 'rime', will refer here specifically to print, to the spelling chunk in a monosyllabic word that corresponds to the rhyming part heard in speech (the letters 'at' comprise the rime unit in the word 'cat'). Thus, a 'rime analogy strategy' will refer to a reading strategy that entails using the knowledge of how to pronounce a particular rime-unit spelling pattern within a word to aid in the pronunciation of a new word that contains that rime unit. The term 'onset' will refer to the initial letter-sound(s) in a word (i.e. /c/ in the word 'cat', /st/ in the word 'stop').

The main focus of this paper is to examine studies which are concerned with rhyme awareness, but causal evidence from rhyme awareness training studies, needed in order to substantiate Claim 2, is limited. Although the ability to hear rhyme in

spoken words and the ability to use a rime analogy strategy to read *printed words* involve very different skills and according to some evidence are not related (Duncan et al, 1997; 2000), others cite evidence from their research that the two skills may be linked (Goswami, 1999a). Yet, despite continuing debate in this area, rime units and analogy do feature as elements in beginning reading instruction. Thus, because of the shortage of rhyme training studies and the perhaps premature influence of rime on instructional practice, it was decided to include evaluation of the additional rime-training study evidence as well.

For Claim 1 and 2 sections, studies are evaluated in tables according to a number of internal and external validity criteria, the overall percentage of studies meeting each criterion appearing across the foot of each table. For the Claim 3 section, a less detailed analysis of the related research is presented.

Finally, the implications for classroom teaching practice are discussed.

Research procedures and selection criteria

Four databases were searched to identify possible studies for inclusion in this review: Psychological Literature Abstracts (PsycLit), the Institute for Scientific Information Web of Science (ISIWoS), the National Library of Medicine (PubMed), and the Educational Resources Information Centre (ERIC). The searches of PubMed and ERIC covered the years from 1966 to the present, while the searches of PsycLit and ISIWoS covered the years from 1983, and from 1981 to the present, respectively. Descriptors used were: 'rhyme/rime and reading', 'rhyme/rime and instruction', and 'rhyme and phoneme awareness' for all databases. Additional studies were located through consulting the reference lists in articles, and frequently cited journals were further hand-searched to ensure all recent articles on the subject were included. The total number of studies arising out of these searches was 415. This number was reduced to 61 after ruling out overlapping studies, those studies obviously off topic, and after applying the following selection criteria.

1) The study had to have been published in a refereed journal. However, as there are few training studies, an exception was made for studies of this type; thus, included were two texts, one in which chapters provide supplementary description to a published article (Bradley & Bryant, 1985), and one where a study not yet published is described in a single chapter (Duncan & Seymour, 2000a); in addition, a thesis, a description of which is due for publication, was included (Watson, 1999).
2) The study had to have been quantitative, investigating one of the specific topics of this paper.
3) Simple correlational studies (such as Majsterek & Ellenwood, 1995) that did not consider the contribution made by other factors to reading, in addition to rhyme, were omitted.
4) The dependent variable of all but Claim-3 studies had to have been some form of reading attainment. In one study omitted, reading was made an independent variable (Maclean et al, 1987). In another study omitted, the dependent variable reading measure consisted of pseudo-words that contained spelling sequences and rimes that do not occur in real words (Yopp, 1988); in contrast, a study by Duncan and others (1997) was included in the analyses because the non-words of the dependent variable reading measure were derived from real words.

5) If two studies reported on the same data and the same sample, with final follow-up test data at the same age, the better conducted of the two studies was selected (Bryant et al, 1990a was examined in preference to Bryant et al, 1990b). Similarly, within studies, some statistical analyses were examined in preference to others; where this occurs, a page number or sample is specified in the tables.

Several training studies were excluded when four further selection criteria were applied (see Appendix A).

1) Studies which lack a control group that receives teaching of some kind were excluded. Those studies which have an unseen or no-training control group do, in effect, simply compare instruction with no instruction. Training studies in which the experimental design is such that subjects function as their own controls were not included either, as without a taught control group, no conclusions can be drawn about the effectiveness of a particular training. All that it is possible to establish is that some form of training is more effective than no specific training at all.
2) As was the case for correlational studies, training studies that do not include some reading measure as the dependent variable were omitted.
3) Training studies that compare rhyme (or rime) training which includes alphabetic instruction with a type of training which does not, were omitted on the grounds that such comparisons represent an inadequately pure test of the effects of rhyme (or rime) instruction. In such cases, it is difficult to determine whether it is the rhyme or the alphabetic training that produces effects on reading.
4) Studies that rely on in-house reading measures exclusively (many such measures are often too short, too narrowly focused, or biased towards one treatment group) were also excluded. In such studies it is impossible to judge whether a particular rhyme/rime training reliably promotes reading ability.

After this culling process, the 61 studies identified as appropriate for investigation were sorted into three categories; those that were concerned with: Claim 1, the correlation or predictive relationship between rhyme and reading (32 investigations within 30 studies); Claim 2, the possible causal links between rhyme (or rime) awareness and reading (13 studies); and finally, Claim 3, the possible connection between rhyme and phoneme awareness (16 studies, 12 analysed in detail). Each category was then further broken down into those studies finding either evidence for, or against each claim.

A total of twelve studies from the three topic areas were randomly selected for rating by a second reader to establish the reliability of the author's coding of studies according to the methodological criteria (6 + 4 + 2 studies for Claim 1, 2 and 3, respectively). A trained graduate student experienced in research methodology independently rated this sub-sample of twelve studies using the appropriate evaluative criteria. There was 97.6% agreement ($\kappa = .952$) across the criteria examined; all coding discrepancies were checked and resolved through re-examination and discussion.

Evaluative criteria

The criteria used to evaluate the methodological quality of studies are defined in Tables 1 and 2. These are based on the work of a number of authors (Campbell &

Stanley, 1966; Spatz & Johnson, 1984; Schult, 1999; Tabachnick & Fidell, 1996, 2001; Wilkinson et al, 1999). The weighting given to each criterion was determined as follows: in cases where failure to satisfy a criterion was undesirable but unlikely to compromise either the internal or external validity of the study, a weighting factor of 1 was assigned. Where failure to satisfy an internal validity criterion placed the conclusions drawn in doubt, a weighting factor of 2 was given. Those criteria regarded as imperative for eliminating rival interpretations and for permitting generalisation of findings to other populations or contexts were allocated a weighting factor of 3.

Table 1 defines the criteria used to evaluate the selected correlational studies. Additionally, there were specific rules established to objectively judge whether a particular Table 1 criterion was met, and important reasons why each criterion should be met. These are now outlined.

Sample, tests, reliability (r), guessing. A full description of the *sample* selected is necessary for replication purposes, and for determining the extent to which findings might apply. If the sample consists of pre-readers, the findings will have stronger implications for beginning reading instruction than a sample of older, reading subjects. An indication of the size of the sample provides some index of how powerful the findings are; the larger the sample, the greater the likelihood that the results reflect what would be true for a much larger population. Description of any in-house *tests* is vital for generalisation and replication purposes, while appropriate test selection is crucial in terms of establishing validity. Although not usually fatal to the validity of a study, the failure to report test *reliabilities* seriously weakens the importance one can attach to the results. Confidence in a test is reduced as the reliability coefficient decreases; (if test–retest reliability is 0.80, the test is capable of obtaining the same results only 64% of the time, whereas if $r = 0.70$, the test is reliable only 49% of the time). Whether or not *guessing* was a factor controlled in these studies was particularly relevant to rhyme and reading research. Several rhyme and phoneme multiple-choice answer tests administered in these studies involve a chance element. If mean scores obtained on such measures are not significantly above a level that would occur simply by chance through random guessing, then the measure cannot be considered an adequate measurement tool. Since statistical analyses conducted using means that are no better than chance levels may yield results that are uninterpretable, the binomial test was applied in order to determine whether mean scores met this criterion with α set at 0.05.

Rhyme only. Any conclusions as to whether rhyme is, or is not, related to reading depends crucially upon the way in which 'rhyme' itself is measured. A potential problem with all tests is the confounding of what is being measured. In this instance, the extent to which a chosen rhyme measure assesses rhyme awareness, and *rhyme only*, is critical in evaluating the validity of research conclusions. Two tests, the sound categorisation oddity test (Bradley & Bryant, 1978; Bradley, 1980) and the rhyme-only oddity version of this test, are particularly problematic in this regard.

The first of these, the sound categorisation test, requires detecting the odd-sounding word out of three or four spoken words. However, the test includes three different conditions, and one score, obtained by combining the results from all three conditions, is used. The three conditions are as follows: a) all the words rhyme and also end in the same *final consonant* except the odd word (bat, cat, hat, *sad*), or b) all

Table 1. Definition of validity criterion used to judge correlational or predictive studies.

Criterion	Definition	Weighting
Sample	Not more than one of the following details is missing: size of sample, type and number of schools, subjects' mean age, gender, race and socio-economic background. If attrition rate is high, remaining sample is checked that it is still representative.	3
Tests	A full description of *all* in-house measures is supplied (instructions, number of items per test, choices per item).	2
r	Test reliabilities of the majority of measures, and in particular all in-house measures, are reported.	2
Guessing	There are checks where necessary (in measures with multiple options) that the means obtained are *significantly* above what could occur through guessing (established through a binomial test where $p < .05$). Where test scores do not meet this requirement, scores are either transformed or not used in the subsequent multiple regression analyses.	3
Rhyme only	The test used to measure rhyme is not contaminated by measuring something else as well. (Alliteration and rhyme measures are not combined to produce the only rhyme score.)	3
Means	All means and standard deviations for both dependent and independent variable data are reported.	2
Distribution	Dependent variable data are normally distributed, except in the case of pre-readers' scores where floor effects are expected.	2
Variances	The homogeneity of variance assumption is satisfied (the spread of scores between all measures is roughly equal. The rule used was that the largest standard deviation (of independent variables) had to be no more than twice as large as the smallest.	2
Skewness	Where the statistical test used is not robust to deviations from normal (multiple regression analysis for example) the independent variable data is examined for effects of skewness and for outliers, and if necessary, especially if N is not large, an accepted method is used to address this problem.	3
Collinearity	A complete correlational analysis of the independent variables is reported. Selection of predictor variables for multiple regression analysis have strong correlations with the dependent variable, but only weak correlations with each other (not more than *r* of 0.50).	2
Ss to variables	The ratio of subjects to independent variables is acceptable. For multiple regression analyses, the ratio accepted was $\geqslant$15 to 1, and for Anova, 10 subjects per cell was accepted unless both Ns and variances were unequal.	3
Extraneous factors:	Type I errors are avoided by taking into account important extraneous factors which might influence reading ability. Subject variation in these factors is statistically controlled or removed.	
1	Age: The effects of subjects' age is controlled via use of age standardised reading scores, or age is entered as a separate variable in regression analysis.	3
2	IQ: Some measure of verbal or non-verbal IQ was administered and subject variation in IQ is partialed out, or controlled as a factor.	3
3	PhA: Subject variation in phoneme awareness ability is removed or controlled as a factor.	3
4	ABC: Effects due to alphabet knowledge in the case of pre-readers (letter names or sounds) or to prior reading ability in the case of readers are statistically controlled or removed.	3
5	Other: In separate analyses, or if subject to variables ratio in regression analysis permits it, the contribution made by some other variable to reading ability is investigated. For example, speech perception or production, speed of processing (rapid naming or speech rate), short-term or working memory.	1
	Maximum score	**40**

the words rhyme and also have the same *medial vowel* except the odd word (bat, cat, hat, *sit*), or c) none of the words rhyme, but all begin with the same *initial consonant sound*, except for the odd word (ham, hat, had, *tap*).

Since condition c requires phoneme detection, and conditions a and b appear to involve rhyme detection, this is an example of a confounded measure. One must ask what exactly is being measured; is it the ability to detect phonemes or rhyme? Although in some studies, this sound categorisation oddity measure is the only test used to assess rhyme, some have interpreted the results from such studies as constituting evidence that rhyme, on its own, is related to reading. One can only legitimately conclude, however, that the awareness of both rhyme and phonemes, if measured in combination, might be related to reading.

In an attempt to avoid this kind of problem, most researchers use the rhyme oddity test, where test items are confined to the two rhyming conditions, conditions a and b above; where this was the case, the studies evaluated here were not penalised for using a confounded rhyme measure. Even with this version of the oddity test, however, the confounding of rhyme and phoneme measurement may still occur. This is because under rhyming condition a, a child could choose the correct answer *either* by noticing that one word does not sound the same somehow overall (it does not rhyme), *or* by noticing that one word does not end in the same final sound (final phoneme detection). Indeed, some evidence demonstrates that performance on the rhyme oddity test is highly correlated with performance on phoneme tasks (correlations as high as 0.67). In contrast, among the studies evaluated here, the highest correlation reported between performance on a rhyme generation test (a purer rhyme measure) and a phoneme test was 0.35.

Other findings fuel further speculation as to whether rhyme or phoneme detection ability is being measured by the rhyme oddity test. Kirtley and others (1989) found that the ability to detect final phonemes is particularly predictive of reading level; parallel to this finding, others have found that, compared to performance on condition b above, performance on condition a (final phoneme differs in the odd word) is more predictive of reading ability. In fact, Carillo (1994) found that scores on the two rhyming conditions, a and b, which both supposedly measure rhyme detection, were not highly correlated with each other. Instead, scores obtained from rhyming condition a (final phoneme different) were most highly correlated with final phoneme segmentation scores. And, condition b scores (medial vowel different) were most highly correlated with total segmentation scores.

Apart from the doubt over whether it is rhyme or phoneme awareness which is being measured by either the sound categorisation or rhyme oddity test, evidence suggests there may be at least four other contaminants operating within these tests. These include: the child's working memory (Oakhill & Kyle, 2000), speech perception (Snowling et al, 1994), and socio-economic background (Burt et al, 1999), as well as the examiner's quality, volume or pace of speech production while articulating the three or four options per item (McDougall et al, 1994).

Means, distribution, variances. Reporting of test *means* is vital in judging the quality of a study, and decisions on whether or not four other criteria are met (*guessing*, *distribution*, *variances* and *skewness*), are necessarily restricted. One of the assumptions of statistical tests is that the dependent variable data are *normally distributed*. However, among young children, reading scores are often at floor. As this was

expected and statistical analysis is quite robust to violations of this assumption, this was accepted. Studies that did not report the reading measure standard deviations but reported normally distributed scores were rated positively. In studies conducting multiple regression analysis, the dependent variable was reading ability; in studies where analysis of variance (Anova) was used, the dependent variable often consisted of other variables that might characterise a particular reader group. If *variances* under comparison are unequal, the chance of reporting a significant difference, when none exists, increases. Thus, although violation of the homogeneity of variances assumption does not invalidate the results from parametric statistics, it does weaken them considerably. The rule used here in deciding whether a study met this criterion or not was that the largest standard deviation had to be no more than twice as large as the smallest. In studies conducting multiple regression analysis, where the contribution of different variables to reading ability was being compared, the standard deviations of all the independent variables entered into the relevant multiple regression equations were examined. In the case of studies using Anova, whether dependent variable variances were equivalent was the focus.

Skewness, collinearity. In studies which conducted multiple regression analyses, the data were examined further for *skewness*, independent variable data ceiling or floor effects, presence of outliers (extreme scores), and for evidence that, where such problems existed, steps had been taken to remedy them. Failure to satisfy the skewness criterion was serious, as it can lead to spurious results, particularly if only some of the scores entered into a multiple regression equation are skewed while others are normally distributed. Skewness was identified in cases where all scores on a measure were either extremely high or low (skewness $\geqslant \pm 2$). Outliers were suspected where a test had a high mean as well as a very high standard deviation (higher than 2 standard deviations). Violation of the *collinearity* criterion is not fatal to the validity of a study's findings, but it can limit interpretation of the results obtained. If two highly-correlated independent variables are entered into a regression equation, it can make interpreting the importance of each of these variables difficult because the effects are confounded due to the correlation between them. A limit of a correlation coefficient of 0.50 between variables was set, so that any variables entered in the same regression equation that were more highly correlated than this were judged to have violated the collinearity assumption.

Subject to variable ratio. In conducting multiple regression analyses, the ratio of *subjects to independent variables* needs to be sufficient or the results obtained will be meaningless. Although a ratio of 10 to 20 subjects per variable has been recommended (Tabachnick & Fidell, 1996), more recently, a stricter rule of thumb is suggested where the minimum number of subjects must be equal to $50 + 8$ times the number of variables (Tabachnick & Fidell, 2001). As a compromise, the ratio accepted in this evaluation was 15 subjects per variable. This ratio is quite lenient, especially in cases where dependent measures are skewed or test measures are unreliable. Where it was not possible to determine the ratio, this criterion was rated negatively. With analyses of variance (Anova), a ratio of 10 per cell was accepted; a ratio lower than this can lead to incorrect or misleading results. In particular, small or unbalanced sample sizes can increase vulnerability to assumption violations.

Extraneous variables. Due to the selection criteria set up, all 32 correlational studies went beyond investigating concurrent or longitudinal association between rhyme and reading and attempted to determine what other variables besides rhyme awareness might relate to reading ability, and how strongly. There were two different ways in which this was done. About two-thirds of the studies used forward stepwise regression analysis which allows one to examine if a relationship still exists between rhyme and reading once the effects of other variables have been removed. The remaining studies examined the extent to which different variables were associated with reading through the use of analysis of covariance, or analysis of variance using matched comparison groups, matched on a number of factors.

It may seem that the use of such methods addresses the non-spuriousness criterion mentioned earlier and makes possible the establishment of causation. However, this is not the case. Although such methods permit one to state that *X* is related to *Y*, or even that *X* 'predicts' *Y* after taking into account the influence of some other factors, they do not allow one to assert that *X* 'causes' *Y*. These statistical methods only help to identify relationships among variables, but strong relationships between variables can stem from many of the sources not measured.

However, factors such as age, intelligence, phoneme awareness and alphabet knowledge are all known to contribute to reading achievement. Without taking these known influences into account while conducting statistical analyses, the risk of committing a Type-I error – concluding that rhyme is significantly related to reading ability when in fact it is not – is seriously increased. Thus, in cases where the subject to variable ratio permitted it, whether these four factors were ruled out as alternative explanations of reading variance before considering the contribution of rhyme awareness was deemed important. If factors other than age, IQ, phoneme awareness and alphabet knowledge were investigated, and the subject to variable ratio permitted it, this too, was deemed useful, but not critical to a study.

Age. One factor known to relate to reading achievement is chronological *age*. Correlations as high as 0.74 have been found between age and reading achievement (Nation & Hulme, 1997). When chronological age varies widely within the sample investigated (a three-year age-range, for example), age can account for as much as 31% of the variance in reading (Kirtley et al, 1989). Furthermore, performance on the rhyme oddity task is highly related to age (Nation & Hulme, 1997; Ackerman & Dykman, 1993). When rhyme is measured in this way, chronological age is also being measured. For these reasons then, when assessing the extent to which rhyme awareness might be related to reading ability, unless age is taken into account in some appropriate manner, it is equally likely to be chronological age, not rhyme awareness, which is responsible for variation in reading scores. To be rated as satisfying the age criterion in this evaluation, a study had to have used age as a separate variable in analysis or the dependent variable reading scores used had to have been age-standardised scores. This criterion is rather strict since standardised IQ and vocabulary scores do control age to some extent. However, the age control is specific to IQ or vocabulary and thus these variables do not control all the variance that can be accounted for by chronological age in general. For the studies where different groups of readers were compared, age was deemed controlled only when groups were matched on reading age.

Intelligence. When examining the relationship between rhyme awareness and reading, the risks of drawing erroneous conclusions are high unless the effects of *intelligence* are partialled out. Since intelligence is a factor that can explain as much as 36% of the variation in reading ability (McDougall et al, 1994), the apparent amount of contribution rhyme makes to reading will be much inflated unless the variance due to intelligence is removed. Indeed, performance on rhyme tasks is found to be highly correlated with intelligence (Ackerman & Dykman, 1993), with correlations found as high as 0.59 (Bryant et al, 1990b), suggesting that part of what rhyme tasks measure is intelligence. Although vocabulary was measured as well as, or instead of, IQ in a number of the studies evaluated here, it is not a useful substitute for intelligence. When standardised age-related vocabulary scores are used in statistical analyses, vocabulary is found not to contribute at all to reading ability (McGuinness et al, 1995), unless there are extremely low scores present in the sample (representing the lowest 9% of the population) (Hurford et al, 1994). Thus, to satisfy this criterion, verbal or non-verbal intelligence scores had to have been used as a separate variable in the multiple regression equation, or as a covariate in analysis of covariance (Ancova); or if different reader groups were being compared, they had to have been matched on intelligence scores.

Phoneme awareness. As already noted, some tests designed to measure rhyme awareness may actually be measuring *phoneme awareness*. For this reason, and also because phoneme awareness is highly correlated with reading ability (correlations as high as 0.80 or more among the studies examined here) when considering whether rhyme relates to reading, the separate variance that may be due to phoneme-awareness ability needs to be taken into account. However, it is important that the particular phoneme tests used are not so difficult that they result in floor effects, a common problem among samples of very young children. When the obtained phoneme awareness scores are very low (severe positive skew) while the rhyme awareness scores are normally distributed the chances are high that the variance in reading due to phoneme awareness ability will not be properly accounted for and instead, it will be attributed to rhyme awareness.

Alphabet knowledge. Since phoneme awareness and alphabet knowledge are highly correlated, and since the phoneme measures used in these studies were often too difficult to produce normal score distributions, ruling out the contribution of *alphabet knowledge* among non-reading children (or prior reading ability among reading children) was considered an imperative as far as the internal validity of a study was concerned. In fact, many of the studies here focused on what factors were predictive of reading ability among pre-school, non-reading children. One obvious factor which might determine subsequent reading ability is initial reading ability, but this skill is often very difficult to assess adequately at this stage. Variation in young children's alphabet knowledge, however, is easy to measure and it has been found to be one of the most important predictors, if not *the* most important predictor, of subsequent reading ability (Duncan et al, 1997; Duncan & Seymour, 2000b; Johnston et al, 1996; Majsterek & Ellenwood, 1995; Muter et al, 1998; Stuart, 1995). Alphabet knowledge is also highly correlated with rhyme; the chances are high that a child who has good rhyme awareness will also have good alphabet knowledge. Thus, if alphabet knowledge is not taken into account among pre- or beginning readers, a

good part of the variance in reading which is attributed to phoneme or rhyme awareness may actually be due to alphabet knowledge.

In concurrent investigations involving older, reading children where prior reading ability was not assessed earlier in time, or in studies involving different reader groups, this criterion was rated as non-applicable. Therefore, this criterion was quite lenient in the case of some studies involving developing readers where letter knowledge could have been assessed.

A different set of methodological criteria were used to evaluate training studies (Table 2). Included in this table are brief explanations of the consequences of failing to satisfy each criterion. The majority of the criteria were used to judge the validity of a study's findings (internal validity), but several criteria were used to evaluate the extent to which the findings from the study might be practically applied. Since rhyme and analogy instruction is an integral part of beginning reading instruction prescribed by the government's *National Literacy Strategy*, one criterion, the *Beginning readers* criterion, is useful in helping to identify where evidence exists for teaching beginners to read by such methods. For the Table 2 criteria that were the same as Table 1, the same judging rules were applied.

Finally it is important to note that good quality reading research involving young children, schools and classrooms is very difficult to conduct. In rating these studies, no indictment against any researchers or particular studies is intended.

Results

Claim 1: rhyme is related to reading

There were 32 relevant investigations identified (in 30 articles) from the search procedures, and these were each evaluated in Tables 3 and 4 according to the criteria defined in Table 1. All the studies in this section had a similar purpose – to determine whether rhyme awareness might be related to either concurrent or subsequent reading ability. Thus, the sixteen evaluative criteria described in Table 1 are especially relevant to correlational studies which seek to establish concurrent association (17 investigations) or longitudinal, predictive association (time order) (15 studies). The studies evaluated in Tables 3 and 4 measured rhyme awareness and reading among preschool non-readers (12 studies), among older reading subjects (11 studies), and between different reading ability groups (9 studies) (see Appendix B). Among the 32 investigations, sixteen found that rhyme was significantly related to reading (Table 3), while sixteen found that it was not (Table 4). A summary of research quality, according to each of the Table 1 methodological criteria, follows:

Sample, tests, reliability, guessing. Sixty-three percent of the studies in Table 3 (finding that rhyme was related to reading) and 81% of the studies in Table 4 (finding that rhyme was not related to reading) met the sample description criterion. The non-standardised, in-house measures used in these studies were described reasonably well in the Table 3 studies (63%) and very well by all of the Table 4 studies (100%). Among both Table 3 and 4 studies, reporting of test reliabilities was not widely practised: only five out of sixteen studies (31%) in both Table 3 and Table 4 studies met this criterion. Among the studies that did report test reliabilities

Table 2. Definition and rationale of methodological criteria used to evaluate training studies.

Criterion	Definition	Weighting
Indicators of study applicability:		
Sample description	The sample selected is described providing at least 5 of the following details: number of subjects and mean age, schools or classes, SES, gender, language	2
Beginning readers	Participants were beginning readers; pre-instructional reading measures indicate most subjects are unable to read.	1
Treatment description	The training given to all comparison groups is described in sufficient detail to allow study replication.	2
Tests	A full description of all in-house measures is supplied (instructions, number of items and choices per item).	2
Indicators of study validity:		
a) Ensuring equivalency between comparison groups		
Random assignment	Subjects are randomly assigned to different instructional conditions to control for individual differences which might determine response to instruction.	3
Age	The mean age for each group is reported and comparison groups do not differ significantly in age.	3
IQ	A standardised intelligence or vocabulary measure was administered, and groups did not significantly differ prior to instruction.	3
Reading age or abc knowledge	Pre-treatment differences in either reading age (for reading subjects) or alphabet knowledge (for non-reading subjects) were equated between groups.	3
b) Ensuring equivalency between training or instruction		
Same materials per group	All groups receiving instruction were exposed to the same materials; groups differed only with respect to instructional *method*.	1
Same time	Instructional time received is reported and was equal for all instructional groups to control for differential effects of teaching time effects	3
Counter-balancing of instructors	Either all groups were taught by the same instructor, or multiple instructors taught all treatment groups: a control for teacher bias.	2

r (test reliabilities)	Test–retest reliability coefficients for all in-house test measures are reported in the article as an indication of the validity of the findings.	2
Instructional time >60 minutes	The amount of instruction given to experimental and comparison training groups is reported and exceeded 60 minutes overall.	2
Adequate N per cell	The number of subjects is sufficient to control probability of Type II error; a minimum of 10 subjects per cell was accepted here.	1
Follow-up effects	Follow-up testing occurred after some delay to check for maintenance effects of the treatment.	2
Absence of ceiling and floor effects	The dependent measures were capable of measuring treatment effects at post-test. (Floor effects are expected at pre-test in young children.)	3
Appropriate unit of analysis	The unit of analysis used corresponded to the unit of treatment, thus satisfying statistical assumptions and minimising the chance of Type I error.	2
Appropriate statistics	Satisfactory statistical procedures were used appropriate to the hypotheses being tested and the purpose of the study.	3
	Maximum	**40**

Table 3. Studies finding that rhyme *is related* to reading cross-referenced with methodological criteria

	Sample	Tests	*r*	Guessing	Rhyme only	Means	Distri-bution	Variances	No skew	Collinearity	Ss to var ratio	Extraneous factors				
												1	2	3	4	5
Weighting:	3	2	2	3	3	2	2	2	3	2	3	3	3	3	3	1
Bowey & Francis, 1991	Y	Y	N	N	Y	N	N	Y	N	N	N	N	(N)	(N)	(N)	(Y)
Bradley & Bryant, 1978	N	N	N	N	N	N	N	N	- (Anova) -		N	n/a*	Y	N	n/a	N
Bradley & Bryant, 1983 (Study 1)	N	N	N	Y	N	N	N	N	N	N	Y	Y	Y	N	N	Y
Bryant et al, 1990a	N	Y	N	N	Y	Y	Y	Y	N	N	N	Y	Y	N	N	Y
Carduso-Martins, 1994 (p. 34 only)	Y	Y	N	N	Y	Y	Y	N	N	N	N	Y	N	(Y)	(N)	(N)
Ellis & Large, 1987	Y	N	N	N	Y	N	N	N	- (Anova) -		N	n/a	Y	N	N	N
Gathercole & Willis, 1991	N	N	N	Y	Y	Y	Y	Y	Y	N	N	Y	N	N	N	Y
Hansen & Bowey, 1994	Y	N	N	Y	N	Y	N	N	Y	N	N	N	Y	Y	n/a	Y
Hoien et al, 1995 (Study 2)	Y	N	Y	Y	Y	Y	N	N	N	N	Y	N	N	Y	n/a	Y
Kirtley et al, 1989 (Exp. 2)	N	Y	N	N	Y	Y	Y	Y	N	N	Y	Y	N	N	n/a	N
Lundberg et al, 1980	N	Y	Y	Y	Y	Y	Y	N	N	N	Y	N	Y	Y	N	Y
McDougall et al 1994 (analyses 1, 2, & 4)	Y	Y	N	Y	Y	N	Y	N	N	Y	N	Y	Y	Y	n/a	(N)
Rohl & Pratt, 1996	Y	Y	Y	N	N	Y	Y	Y	N	N	N	Y	N	Y	N	(Y)
Stainthorp & Hughes, 1980	Y	Y	N	N	Y	Y	Y	N	- (Anova) -		Y	n/a	N	N	n/a	N
Wagner et al, 1997	Y	Y	Y	N	N	Y	Y	N	N	Y	Y	N	N	N	N	Y
Wood & Terrell, 1998 (Part II)	Y	Y	Y	N	Y	Y	Y	N	N	N	N	N	N	Y	N	(Y)
Ratio of studies meeting this criterion:	10/16	10/16	5/16	6/16	11/16	11/16	10/16	5/16	2/13	2/13	6/16	7/13	7/15	6/14	0/8	7/10
Percentage:	63	63	31	38	69	69	63	31	15	15	38	54	47	43	0	70

* Different chronological age, good and poor reader groups (matched on reading age), differed in rhyme awareness.
Different reader groups, of same chronological age, differed in rhyme awareness.

Average percentage of criteria met by all studies: 44%

Table 4. Studies finding that rhyme *is not related* to reading cross-referenced with methodological criteria.

	Sample	Tests	*r*	Guessing	Rhyme only	Means	Distri-bution	Variances	No skew	Collinearity	Ss to var ratio	Extraneous factors				
												1	2	3	4	5
Weighting:	3	2	2	3	3	2	2	2	3	2	3	3	3	3	3	1
Carduso-Martins, 1994 (pp. 32–33 only)	Y	Y	N	N	Y	Y	Y	N	- (Anova) -		N	n/a +	N	N	n/a	N
Carillo, 1994 (p. 292, grade 1 only)	Y	Y	N	Y	Y	Y	Y	Y	- (Anova) -		Y	n/a ±	N	N	n/a	N
Cronin & Carver, 1998 (p. 456)	Y	Y	N	Y	Y	Y	Y	Y	- (Anova) -		N	n/a +	N	Y	n/a	N
DeJong & van der Leij, 1999 (p. 464)	Y	Y	N	Y	Y	N	Y	Y	Y	Y	Y	N	Y	N	Y	Y
Duncan & Johnston, 1999	N	Y	N	Y	Y	Y	Y	N	- (Ancova) -		Y	n/a*	Y	Y	n/a	Y
Duncan et al, 1997 (Exp. 1)	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	Y	Y
Hatcher & Hulme, 1999	Y	Y	N	Y	Y	Y	Y	N	N	Y	Y	N	Y	Y	Y	Y
Johnston & Rugg, 1989	N	Y	N	Y	Y	Y	Y	Y	- (Anova) -		Y	n/a*	Y	N	n/a	Y
Johnston et al, 1996	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	N	N	Y	Y	Y
McDougall et al, 1994 (analysis 3)	Y	Y	N	Y	Y	N	Y	N	N	Y	N	Y	Y	Y	n/a	(Y)
Muter & Snowling, 1998	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	N	N	Y	N	(Y)	(Y)
Muter et al, 1997 (T2)	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	N	N	Y	Y	(Y)	(N)
Naslund & Schneider, 1996	N	Y	N	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y
Nation & Hulme, 1997	Y	Y	Y	Y	Y	Y	Y	Y	N	N	Y	Y	N	Y	N	Y
Stanovich et al, 1984	Y	Y	Y	Y	Y	Y	Y	N	N	N	Y	N	Y	Y	(Y)	(N)
Stuart, 1999	Y	Y	N	Y	Y	Y	Y	Y	- (Anova) -		Y	n/a ±	N	Y	Y	Y
Ratio of studies meeting this criterion:	13/16	16/16	5/16	15/16	16/16	14/16	15/16	9/16	5/10	8/10	11/16	2/10	9/16	10/16	6/8	9/12
Percentage:	81	100	31	94	100	88	94	56	50	80	69	20	56	63	75	75

* Different chronological age, good and poor reader groups (matched on reading age), *did not* differ in rhyme awareness.
\+ Different reading ability and chronological age between comparison groups, but *did not* differ in rhyme awareness.
± Different reading ability, same chronological age, but *did not* differ in rhyme awareness.
Average percentage of criteria met by all studies: 73%

and where no ceiling effects were operating, the rhyme tasks used had reliability coefficients ranging from 0.65 (a rhyme choice task) to 0.92 (age 6+ rhyme detection task). The reliabilities reported for the phoneme tasks used were generally higher, ranging from 0.78 (Rosner phoneme deletion task) to 0.97 (a phoneme deletion task). All the reliability co-efficients reported for any letter knowledge tests were equal to, or greater than 0.94. As far as the guessing criterion was concerned, only 38% of Table 3 studies in contrast to 94% of Table 4 studies, met this criterion.

Rhyme only. Among the 32 studies reviewed in this section, there were five different types of tests used to measure rhyme awareness. Three studies administered two different rhyme measures and treated the scores separately in subsequent statistical analyses (Bryant et al, 1990a; Muter et al, 1997; Stanovich et al, 1984). The tests involved the following different kinds of tasks, and the number of studies using this particular task to measure rhyme ability is shown in brackets:

1) Bradley and Bryant's sound categorisation task (6 studies);
2) the rhyme only oddity task (15 studies);
3) a rhyme generation or production task (6 studies);
4) a target word or picture/ multiple choice answer task (5 studies);
5) a yes/no task with rhyming pairs of words or pictures (3 studies).

Scores obtained from the Bradley & Bryant sound categorisation test (a confounded measure) were the *only* rhyme measure scores relied upon in the statistical analyses of five Table 3 studies (Bradley & Bryant, 1978; 1983; Hansen & Bowey, 1994; Rohl & Pratt, 1996; Wagner et al,1997). The majority of the all Table 3 and 4 studies used a variant of this test, the rhyme oddity test, where test items were confined to the two rhyming conditions only. Thus, only 35% of Table 3 studies failed to meet the rhyme-only criterion. However, if the judging criteria had been stricter, and had required rating both the sound categorisation *and* rhyme oddity tests as representing confounded measures, a much higher incidence of invalid rhyme assessment would have been found. Under these stricter criteria, one would have had to conclude that most of the studies (81%) claiming that rhyme was related to reading had used a confounded rhyme measure, while almost half the studies (44%) finding the opposite, had also done so.

Means, distribution, variances. Most of the studies listed in Tables 3 and 4 reported all test means and standard deviations, 69% and 88% of studies, respectively. In the majority of studies, dependent variable scores were rated as normally distributed (for Table 3 and 4 studies, 63% and 94%, respectively), although, in a number of cases, floor effects were obtained on pre-readers' reading scores. The equal variances criterion was met by only 31% and 56% of Tables 3 and 4 studies, respectively.

Skewness, collinearity. The skewness criterion was satisfied by only 15% and 50% of Table 3 and 4 studies, respectively, while the collinearity criterion was met by only two of thirteen studies (15%) in Table 3, and eight of ten studies (80%) in Table 4.

Subjects to variables/cells ratio. Even though rather lenient evaluative criteria were permitted, only 38% of Table 2, and 69% of Table 3 studies were rated positively on this criterion.

Age. Age was a factor not very well controlled particularly in Table 4 studies, with only 54% and 20% of Table 3 and 4 studies meeting the criterion, respectively.

Intelligence. Confirming that vocabulary scores are not a useful substitute for intelligence scores, the average amount of variation in reading accounted for by vocabulary scores versus verbal and non-verbal intelligence scores in the 32 investigations examined here was 1.3% versus 19.2%. Seven out of fifteen Table 3 studies (47%), and nine out of sixteen Table 4 studies (56%) controlled intelligence satisfactorily according to the criteria outlined.

Phoneme awareness. In twelve out of sixteen Table 3, and three out of sixteen Table 4 studies that conducted multiple regression analysis, the phoneme scores were positively skewed while the scores obtained on the rhyme measure were normally distributed. For the Table 3 studies, this situation may have meant that rhyme awareness was only found to be related to reading because the variance due to phoneme awareness was not adequately removed; for the Table 4 studies in this same category, it suggests that even when the chance of finding a relationship between rhyme and reading is enhanced by the nature of the data and its analysis, a relationship does not exist.

Far less problematic was a situation that occurred in one Table 3 and one Table 4 study. In the first study, *both* rhyme and phoneme scores were negatively skewed (Hoien et al, 1995), while in the second, *both* rhyme and phoneme scores were positively skewed (Johnston et al, 1996). Unlike the situation described above, in cases such as these where two variables are skewed in the same direction, the over-estimating of variance due to one variable at the expense of under-estimating that due to the other variable does not generally occur.

As it is widely known that phoneme awareness is related to reading ability, almost two thirds of researchers conducting the Table 4 studies attempted to take this factor into account in their analyses (63%). Furthermore, all of these studies found that while rhyme was not related to, or a predictor of reading ability, phoneme awareness was. In contrast, among the fourteen Table 3 studies where control of this variable was appropriate, less than half the researchers (43%) considered it a factor important enough to control. In all the remaining Table 3 studies, where the contribution of phoneme awareness was investigated, reading was found to be more strongly associated with phoneme than rhyme task performance.

Alphabet knowledge. Even though researchers concede that letter knowledge influences reading achievement, they may fail to take this factor into account in their statistical analyses (Bryant et al, 1990a). In fact, none of the eight studies of Table 3 where younger children were involved and where the subject to variable ratio would have permitted it, actually measured and ruled out the effects of prior reading ability or alphabet knowledge. In contrast, among the eight Table 4 studies where control of these factors was appropriate, six out of eight studies did consider their influence.

Other extraneous variables. Several studies investigated the possible relationship between reading and other variables. Some of the following factors were found to

make a significant contribution to reading: speech rate (McDougall et al, 1994), expressive language and receptive language (Bryant et al, 1990b), auditory working memory (Gathercole & Willis, 1991; Hansen & Bowey, 1994) and sentence segmentation (Wood & Terrell, 1998). In finding these results, however, all five of these studies failed to first take into account at least one of the designated four, important variables (age, IQ, phoneme awareness or alphabet knowledge/reading ability). Where subject to variable ratios permitted, 70% of Table 2 and 67% of Table 3 studies investigated the relative contribution to reading of other variables. (Any extraneous factors rated with a 'Y' or a 'N' in brackets indicates that the subject to variable ratio did not permit including this additional variable in the statistical analyses, and when tabulating the percentage of studies that satisfied each criterion these items were not counted.)

Summary

Although, the overall methodological rigour of the studies finding no connection between rhyme awareness and reading was superior to that of the studies finding the opposite (on average, a Table 4 study met 73% of the criteria whereas a Table 3 study met 44%), it may be more meaningful to examine the degree to which just eight of the more critical, heavily weighted criteria were met by all studies. Of the sixteen evaluative criteria in the tables, there are just eight criteria where a failure to meet one of them would seriously threaten study validity: *rhyme only*, *guessing*, *no skew*, *subject to variable ratio*, control of: *age*, *IQ*, *phoneme awareness* and *prior reading ability* or *alphabet knowledge*.

The *rhyme only* criteria may be the most important of these. Studies which base their results on scores which include the measurement of rhyme and something else as well, make any conclusions about the connection between rhyme on its own, and reading impossible. The next three criteria, *guessing*, *no skew* and the *subject to variable ratio*, must also be met if one is to have any confidence in the conclusions of a study. After eliminating all the studies in both Tables 3 and 4 that failed to meet these criteria, no Table 3 studies remain. On the other hand, there remain seven Table 4 studies. Of particular relevance and importance, four of these involved beginners and were longitudinal, investigating what factors in beginning, or non-reading children predict reading ability later on (DeJong & Van Der Leij, 1999; Duncan et al, 1997; Naslund & Schneider, 1996; Stuart, 1999).

So far, then, this exercise shows that rhyme is only found to be related to reading when the tests used are problematic (producing inadequate data sets where there are confounded rhyme scores, scores no different than would be obtained through chance or score distributions that are severely skewed), or when the subject to variable ratio is insufficient to support statistical analysis.

Further inspection of how well extraneous factors (the remaining heavily weighted criteria) were controlled among all studies shows that alphabet knowledge, among Table 3 studies, and age, among Table 4 studies, was the worst met criterion. Thus, a connection between rhyme and reading may only be found when prior reading ability or knowledge of the alphabet is ignored as a factor that may contribute to the variance in reading. However, even when age is not properly controlled as a factor (most Table 4 studies), there is still no relationship found between rhyme and reading.

On balance, considering the results of this evaluation as a whole, the evidence is unconvincing that a relationship exists between rhyme and reading. The main points which support this conclusion are:

1) based on the large discrepancy between Table 3 and 4 studies in the degree of methodological rigour upheld, the evidence in support of no relationship between rhyme awareness and reading is more convincing than the evidence in support of the opposite claim;
2) approximately one-third of Table 3 studies (studies frequently cited in the research literature in support of the claim that rhyme is related to reading) used invalid rhyme measures, effectively rendering their results invalid in the context discussed here;
3) examination of the data sets used in Table 3, compared to Table 4, studies reveals that floor effects on phoneme measures likely resulted in the over-estimation of the importance of rhyme, at the expense of under-estimating the importance of phoneme awareness, to reading;
4) when the subject to variable (or cell) ratio in a study is too low or unbalanced, the results obtained can be uselessly incorrect or uninterpretable; this was the case in six out of eleven Table 3 studies that used an acceptable rhyme measure;
5) where it would have been appropriate, none of the Table 3 studies took alphabet knowledge or prior reading ability into account; when these factors *are* taken into account, no connection is found between rhyme and reading.

On the issue of whether time order has been established by these studies (see Appendix B), among the fifteen longitudinal studies conducted, six found that early rhyme awareness was related to subsequent reading ability, while nine studies found that it was not. However, the quality of the evidence, in methodological terms, from studies in the former group is inferior to that of studies in the latter group; the average percentage of methodological criteria met by these two groups of longitudinal studies was 41% versus 75%. In addition, the number of studies that measured rhyme among pre-readers or early beginners was only five among the first group, in contrast to eight among the second group. Thus, this separate examination of the studies investigating time order helps in confirming the overall conclusion: evidence to date does not support the idea that early rhyme awareness is importantly related to reading ability.

Claim 2: rhyme awareness determines reading ability

Based on the present evaluation of correlational studies, evidence to date insufficiently supports the notion that rhyme awareness and reading ability are associated, and since *both* correlational *and* training-study evidence are needed in order to establish causality, one could argue that there is little point in examining further evidence. However, if the evidence from training studies were found to be convincing, perhaps the correlational research on this topic might warrant further pursuit. Accordingly, after applying the selection criteria outlined earlier, thirteen studies were identified for evaluation.

Among the thirteen training studies that could be identified, only four found that rhyme or rime training produced positive effects on reading, while ten found that such training was either similarly or less productive than other forms of teaching. As

Table 5. Studies finding that rhyme or rime-related training *does* lead to improved reading achievement

Type of experimental treatment* and study	Sample	Beginning readers	Treatment description	Tests	Random assignment	Effects taken into account		
						Age	IQ	R.A. or abc
Weighting: →	2	1	2	2	3	3	3	3
2 Rhyme awareness + phoneme/syllable (NO L-S$^+$)								
Lundberg, Frost & Peterson (1988)	Y	Y	Y	Y	N	N	Y	Y
3 Rhyme awareness + rime units + /or analogy + L-S:								
Santa & Hoien (1999)	Y	Y	Y	Y	N	Y	N	Y
4 Rime unit instructions only:								
Greaney & Tunmer (1996) Exp. 2	Y	N	Y	Y	N	Y	N	Y
Greaney et al (1997)	Y	N	Y	Y	Y	N	N	N
Ratio of studies meeting criterion:	4/4	2/4	4/4	4/4	1/4	2/4	1/4	3/4
Percentage:	100	50	100	100	25	50	25	75

On average, % criteria met: 64%

* Study type (type of instruction received by Group A in each study):
1 = Rhyme awareness training alone; 2 = Rhyme awareness + some phoneme and/or syllables (no L-S); 3 = Rhyme $^+$ L-S = Letter-sound instruction

shown in Table 5, 6 and 7, the experimental training of primary interest within these studies fell into one of four basic types. While a number of studies compared more than two instructional treatments, as shown in Table 7 the differential effects produced as a result of training given to groups A and B are of primary interest, as these were the two groups who received the most similar types of instruction.

Summary

The average percentage of Table 2 methodological criteria met by Table 5 and 6 studies did not differ substantially, 64% versus 71%. Where there was a more important difference, however, was in the nature of the studies conducted: whether there were studies that truly investigated causation between rhyme and reading, where the findings applied to beginning reading instruction, and where the comparisons of instructional treatments were fair.

Only one study (Table 6) could be identified that did not combine rhyme awareness training with something else (Duncan and Seymour, 2000a). This longitudinal study was directed at investigating the effects of pre-school rhyme awareness training *alone* (Study Type 1), on subsequent reading ability one and two years later. After five months of training in nursery school, the rhyme group was found to have significantly superior rhyming skills to those of a group trained in initial phoneme awareness. However, the superior rhyme skills of the experimental group did not appear to have had an effect on reading achievement. No significant differences were found between the two groups in their British Ability Scales (BAS) reading scores later on, in either year one or two of school.

cross-referenced with methodological criteria.

Same materials	Same time	Counter-balancing	r	Instuctional time > 60 mins	Adequate n per cell	Follow-up effects	Absence of ceiling and floor effects	Appro-priate unit of analysis	Appro-priate statistics	
1	3	2	2	2	1	2	3	2	3	Max. 40
N	N	N	Y	Y	Y	Y	Y	N	Y	
N	N	N	Y	Y	Y	Y	Y	N	Y	
Y	N	N	N	N	Y	N	Y	Y	Y	
Y	Y	Y	N	Y	Y	Y	Y	Y	Y	
2/4 50	1/4 25	1/4 25	2/4 50	3/4 75	4/4 100	3/4 75	4/4 100	2/4 50	4/4 100	

awareness + rime unit and/or analogy + L-S; 4 = Rime unit instruction

The remaining twelve studies are concerned with investigating either the effects of combined rhyme and phoneme awareness training or the effects of training involving rime units. None of these studies are useful in terms of investigating a causal connection between rhyme and reading because either they involve confounded rhyme awareness training where it is impossible to tell which element or combination of elements within the training might have had positive effects on reading ability, or they involve teaching rime unit or analogy skills, skills that are markedly different from, and maybe wholly unrelated to, the awareness of rhyme.

Furthermore, setting aside the debate over whether rhyme awareness and rime use are related, those studies that involved rime instruction were only found to produce positive effects among older, already-reading children. Amongst beginning non-readers, this sort of instruction did not enhance reading ability over other forms of instruction; in fact, other forms of instruction produced significantly superior reading progress among children of this age (Deavers et al, 2000; Solity et al, 1999; Watson, 1999).

Because of the initial selection procedures set up in Appendix A, none of the studies selected for evaluation had comparison groups designated Group A or B in Table 7 where one group received letter-sound instruction and the other did not. However, there were still some unfair comparisons made between groups among the Table 5 studies. In one study, the children that received the joint rhyme/rime instruction were individually tutored, while the comparison children received instruction in small groups (Santa & Hoien, 1999). It is impossible to tell, in this case, whether the positive effects found for the rhyme/rime group were due to the individual tutoring or to the type of instruction. Two other Table 5 studies may have

Table 6. Studies finding the rhyme or rime-related training does *not* lead to improved reading achievement

Type of experimental treatment* and study	Sample	Beginning readers	Treatment description	Tests	Random assignment	Effects taken into account		
						Age	IQ	R.A. or abc
Weighting: →	2	1	2	2	3	3	3	3
1 Rhyme awareness alone:								
Duncan & Seymour (2000a)	Y	Y	Y	Y	N	Y	Y	Y
2 Rhyme awareness + phonemes +/or syllables (NO L-S^{+})								
Barron et al (1992)	Y	Y	Y	Y	Y	Y	Y	Y
Bradley & Bryant (1983)	Y	Y	Y	N	N	Y	Y	N
3 Rhyme awareness + rime units +/or analogy + L-S								
Deavers et al (2000)	Y	N	Y	Y	N	Y	N	Y
O'Shaughnessy & Swanson (2000)	Y	N	Y	Y	Y	Y	Y	Y
Qi & O'Connor (2000)	Y	Y	Y	Y	Y	N	Y	Y
Solity et al (1999)	Y	Y	Y	Y	N	Y	N	Y
Watson (1998)	Y	Y	Y	Y	N	Y	Y	Y
4. Rime unit instructions								
Joseph (2000)	Y	N	Y	Y	Y	N	N	Y
Ratio of studies meeting criterion	9/9	6/9	9/9	8/9	4/9	7/9	6/9	8/9
Percentage	100	67	100	89	44	78	67	89

On average, % criteria met: 64%

* Study type (type of instruction received by Group A in each study):
1 = Rhyme awareness training alone; 2 = Rhyme awareness + some phoneme and/or syllables (no L-S); 3 = Rhyme
$^{+}$ L-S = Letter-sound instruction

mounted unfair comparisons as well, since the experimental groups received considerably more letter-sound instruction than the controls (Greaney & Tunmer, 1996; Greaney et al, 1997). And, in the second study, the justification for excluding studies that rely exclusively on the results from in-house reading measures is highlighted. Three out of four post-training tests were in-house measures that required reading isolated words in lists, words that could not be guessed from context, a procedure that placed the control children (taught how to guess words from context) at a distinct disadvantage.

Finally, unlike all but four of the 13 training studies examined here, the correlational evidence examined in the Claim 1 section of this paper is concerned with implicit rhyme awareness, an awareness that does not require the explicit or conscious segmentation in words of initial phonemes from rhyming segments. In contrast, nine of the 13 training studies examined in this section were concerned with the explicit awareness of rhyme requiring identification of the rhyming segment within spoken words and the identification of rime units within written words. Since there is some question as to whether these two kinds of awareness constitute the same thing, or are even related, the combination of both correlational and training study evidence on the *same* topic, required in order to establish causality, appears to be lacking.

cross-referenced with methodological criteria.

Same materials	Same time	Counter-balancing	r	Instuctional time > 60 mins	Adequate n per cell	Follow-up effects	Absence of ceiling and floor effects	Appro-priate unit of analysis	Appro-priate statistics	
1	3	2	2	2	1	2	3	2	3	Max. 40
Y	Y	N	N	Y	Y	Y	Y	N	Y	
Y	Y	Y	Y	Y	Y	N	N	Y	Y	
Y	Y	Y	N	Y	Y	N	Y	Y	Y	
N	N	N	N	Y	Y	N	Y	N	Y	
N	Y	Y	N	Y	Y	N	Y	N	Y	
Y	Y	N	N	Y	Y	N	Y	Y	Y	
N	N	N	N	Y	Y	N	Y	N	Y	
N	Y	N	N	Y	Y	Y	Y	N	Y	
Y	Y	N	N	Y	Y	N	Y	N	Y	
5/9	3/9	1/9	9/9	2/9	9/9	2/9	8/9	3/9	9/9	
56	33	11	100	22	100	22	89	33	100	

awareness + rime unit and/or analogy + L-S; 4 = Rime unit instruction

Thus, to review so far, in order to establish a causal connection between rhyme and reading it is necessary to demonstrate: association, time order and non-spuriousness. An evaluation of the research methodology employed by those studies investigating the topic, however, reveals that not one of these three criteria has been adequately met. The citing by Goswami of 19 studies in support of the 'early availability of rhyme and its causal links with reading' (1999, p. 219) appears, in part, to contradict this conclusion. However, as far as constituting a 'causal link' between rhyme and reading is concerned, this assertion is not supported by the studies cited, as an analysis of them shows (see Appendix C).

Claim 3: rhyme awareness leads to phoneme awareness

In attempting to establish a causal connection between rhyme and reading, it is regarded as sufficient to demonstrate association, time order and non-spuriousness; however, if the *mechanism* by which rhyme determines reading ability can also be demonstrated the evidence supporting causality is considerably strengthened. Hence, a third claim offering an explanation of why rhyme might affect or determine reading ability has been made as part of Goswami and Bryant's developmental theory (1990): rhyme awareness leads to phoneme awareness, an ability which, in turn, affects

Table 7. Types of instruction compared in rhyme and/or rime training studies and post-training results.

Study type*	Study	Age	Selected instructional conditions	Relevant post-training results
Table 5 studies:				
2	Lundberg, Frost & Petersen (1988)	6:0 NR	A) Rhyme & syllable awareness, syllable blending, phoneme identity, no print B) Social and aesthetic development skills	Group A significantly outperformed B on word reading 2 years later, at end of grade 2, (but not one year later at end of grade 1)
3	Santa & Hoien (1999)	6:7	A) Individual tutoring in: real books, relating rhyme to rime, letter formation, letter names, onsets, sight words, word sorting on basis of medial vowels/use of context B) Small group instruction in: guided and choral reading and re-reading, no word study (reminder only about letters, vowel patterns, or context)	Group A better than B in Woodcock word recognition, prose and nonword reading on both immediate (end of grade 1) and delayed (end of grade 2) post-tests
4	Greaney & Tunmer (1996) (Exp. 2)	10-13	A) Reading by analogy B) Context cue usage (minimal attention to letters)	After 50 minutes' training Group A significantly better than B in prose-reading accuracy
4	Greaney et al (1997)	8:2	A) Rime spelling instruction supplementing reading recovery type instruction B) Context cue usage supplementing reading recovery type instruction; (attn. to letters rarely)	After 30 minutes' training, and one year later, group A significantly better than B in Burt word reading, and reading and spelling words with common rime units
Table 6 studies:				
1	Duncan & Seymour (2000a)	4:2 NR (non-readers)	A) Rhyme awareness training B) Alliteration awareness (treated control; teacher absences) C) Contaminated control (teachers introduced some rhyme activities)	Both A and C significantly better on rhyme production after the 4 mths' training; no significant differences between groups on BAS word reading at end of one (P1) and two years (P2)
2	Barron et al (1992)	6:1 NR	A) Speech only (onset, rhyme, final phoneme, medial phoneme oddity tasks) without print B) Semantic categorisation control without print	After 3 months' training, no significant differences between all 3 groups on Woodcock word and nonword reading
2	Bradley & Bryant (1983)	6:1 NR	A) Rhyme and onset awareness/ no letters B) Concept categorisation/ no letters	No significant differences on Neale or Schonell reading 2–3 years after individual instructional sessions

3	Deavers et al (2000)	4:4	A) NLP: L-S (letter-sounds), segmenting and blending onset and rime, rhyme awareness B) ERR: L-S , seg. & blending phonemes C) Unseen control: L-S, rhyme, rime, ph. identity	After 2 years of instruction, no significant differences between groups in BAS word reading; B outperformed A in reading of regular and irregular nonwords
3	O'Shaughnessy & Swanson (2000)	7:8 (gr. 2 poor readers)	A) Word analogy training (rhyme awareness, whole word and rime recognition, reading by rime analogy strategy B) Phonological awareness training (rhyme and phoneme awareness, segmenting and blending phonemes with and without print) C) Mathematics training	After 6 weeks' instruction, no significant differences between groups A and B but both outperformed C in Woodcock nonword reading
3	Qi & O'Connor (2000)	5:6 NR	A) Sound categorisation (onset and rhyme /rime identity with and without print B) Phoneme training (blending and segmenting of phonemes with and without print	After 10 weeks' training at end of kindergarten, no significant differences between groups in Woodcock word reading and writing letters and words
3	Solity et al. (1999)	5:4 NR	A) NLS: rhyme/analogy/onset-rime/ L-S/ letter names/sight words/ consonant blends B) ERR: ph. blending/segmenting/ L-S/sight words	Group B outperformed A on all 17 reading, spelling, phonological measures except rhyme awareness oddity, after one year
3	Watson (1999) Exp. 1	5:0 NR	A) Analytic + PA: 16 L-S, onset identity, cvc sight words, phoneme and rhyme awareness B) Analytic: 16 L-S, onset identity, cvc sight words C) Synthetic: 40+ L-S, reading (sounding & blending) of cvc words; spelling (segmenting & blending) of cvc words	After 16 weeks' training, group C sig. better than A and B on: BAS, Clay and nonword reading, Schonell spelling, L-S, and ph. seg. , but same as B on rhyme production; group B better than A on nonword reading; A better than B on phoneme segmenting
4	Joseph, (2000)	6+	A) Traditional – reading and writing of words with shared rime units B) Word Box – phonemes, L-S, seg. and blending C) Word Sort – rhyme, rime, phonemes, L-S	After 12 weeks (16¾ hrs.), groups B and C sig. better than A on Woodcock word recognition; B better than A on nonword reading; no significance differences between B and C

* Study type (type of instruction received by Group A in each study):
1 = Rhyme awareness training alone
2 = Rhyme awareness + some phoneme and/or syllables (no L-S)
3 = Rhyme awareness + rime unit and/or analogy + L-S
4 = Rime unit instruction

reading ability. However, since causality between rhyme and reading has not been demonstrated according to the preliminary criteria set out, an in-depth analysis of studies addressing the issue of mechanism does not seem warranted. Instead, what follows is a brief discussion of the research and its methodology relevant to the first part of this claim (the rhyme/phoneme connection), while leaving aside the second part of the claim (the phoneme awareness/reading connection).

A predictive, not necessarily a causal, relationship is implied by the claim that rhyme awareness 'leads to' phoneme awareness. Thus, three longitudinal studies have been cited in support of this proposal. Two of these report on the same sample of children over a one-year (Maclean et al, 1987) and a three-year period (Bryant et al, 1989). Results appear to demonstrate that nursery-rhyme knowledge predicts alliteration detection (involving single phoneme onsets) in the first study, and first and final phoneme detection in the second study. In the third study (Bryant et al, 1990a), detection of rhyme was found to predict phoneme deletion performance. But, even if is assumed that nursery-rhyme knowledge (in the first two studies) equates to rhyme awareness, in all three studies the mean scores obtained on all phoneme measures in question were not above chance levels. Conclusions based on this evidence, therefore, must be regarded as tenuous.

However, the training studies involving younger children, evaluated earlier in this paper, furnish some causal evidence both for and against the proposal that early rhyme-awareness or rimes may be a 'route into phonemes' (Goswami, 1999a, p. 233). Lundberg and others (1988) found that phonological training which included rhyme awareness training significantly improved phoneme segmentation skills over those attained by a non-phonological control group. However, although the methodological quality of this study is fair (65% of criteria met), the research design is problematic. The phonological training in question also included training phoneme awareness, so it is not possible to tell whether it was the rhyme or the phoneme aspect of the training that led to the improvement. Suggesting that it may have been the phoneme aspect of the training that made the difference, Qi and O'Connor (2000) demonstrated that training that taught *only* phoneme skills produced equivalent improvement in phoneme deleting skills to onset-rime training (i.e combined phoneme and rhyme/rime training).

Indeed, findings from a number of the other training studies appear to demonstrate that rhyme training has no effect on phoneme skills. In one study, it was found that phoneme deletion performance only improved among children who were high in letter-sound knowledge, and whose phonological training included print feedback. The phoneme skill of children low in letter-sound knowledge did not improve with the phonological training (rhyme and phoneme) whether or not they received print feedback (Barron et al, 1992). In another study, the effects of three types of instruction were compared after four terms at school: instruction emphasising small units (phonemes), large units (onsets and rimes), and both small and large unit instruction (Deavers et al, 2000). Children in the first group were found to be significantly better at reading nonwords (a measure of phonemic decoding skill) than children in either of the latter two groups, who appeared to be confused by the instruction they had received. In another study, *National Literacy Strategy* training (focusing on rhyme awareness, use of analogy, onset-rime segmentation, some letter-sound and phoneme awareness instruction) was compared over one year to instruction confined primarily to instruction in letter-sound knowledge, and the blending and segmenting of

phonemes in words with and without print (Solity et al, 1999). After taking initial group differences in phoneme awareness into account, it was found that the latter type of instruction resulted in significantly better scores on two phoneme measures. Watson (1999) demonstrated (in Experiment 1) that although the phoneme segmentation skill of a group of children given some phonological training (phoneme and rhyme) significantly improved over a group not given this training, a third group of children, not given phonological training but who were taught many more letter-sounds than the other two groups and within the same period of time, significantly outperformed both groups on a phoneme segmentation test. Finally, Joseph (2000) found that phoneme along with letter-sound training produced significantly better phoneme segmenting and blending skills than rime training.

Much of this evidence suggests that it is letter-sound teaching, not rhyme or rime instruction, which is responsible for producing phoneme awareness (Barron et al, 1992; Joseph, 2000; Solity et al, 1999; Watson, 1998).

Some of the Table-4 correlational studies lend support to this conclusion. Rhyme production ability in pre-readers was found to make no contribution to either first or final phoneme segmentation ability, once the effects of reading or writing the alphabet were removed (Johnston et al, 1996). Phoneme awareness ability was found to be determined, not by good or poor rhyme awareness, but by beginning reader's letter-sound knowledge (Stuart, 1999). Duncan and others (1997) found that rhyme production ability made no contribution to nonword naming once the effects of letter-sound knowledge and letter-sound decoding skills were removed.

The methodological quality of the latterly cited nine Table 4 and 6 studies is good; on average, these studies met 72% of the appropriate methodological criteria.

Other evidence supports these findings. It has been found that children's level of accuracy in detecting phonemes remains low (20% accuracy) until their level of letter-sound knowledge is high (above 80%) (Duncan and Seymour, 2000b). Both novice and non-readers high in letter knowledge are found to be more sensitive to phonemic units (Bowey, 1994).

Summary

The weight of current evidence suggests that rhyme awareness is not related to phoneme awareness, nor does rime unit training appear to affect the development of phoneme awareness. Instead, it appears that alphabetic instruction plays a key role.

Discussion

As noted earlier, the rhyme research evaluated here is relevant to the first of three 'causal connections' in a theory proposed by Goswami and Bryant (1990). In brief, this theory proposes that there are three determinants of reading ability: rhyme awareness, phoneme awareness and spelling. Additionally, two reasons why rhyme awareness might affect reading level are offered: first, rhyme awareness, while having a direct influence on reading ability, also has an indirect influence via its contribution to phoneme awareness; and second, it is linked to a strategy that children appear to have 'available' from the beginning, the ability to make analogies based on rime units.

The first causal connection in the theory, proposing that rhyme awareness affects reading, has had widespread influence on teaching practice. In North America, both rhyme and rime-based teaching have infiltrated classroom practice to some degree, perhaps due in part to Adams' endorsement of onset and rime as a beginning reading method (1994, pp. 306–322). In the UK, the government has introduced the *National Literacy Strategy* in which teachers are required to teach beginners how to read through the use of rhyme awareness and rime analogy instruction.

There are perhaps three reasons why instructional practice has been affected by this first part of the theory: 1) it appears to be supported by evidence, frequently cited, 2) it proposes a logical sequence of development, and 3) it is based upon a number of plausible assumptions.

Instructional implications

In fact, a number of teaching recommendations have been justified on the basis of the cited evidence, the logic of the proposed sequence and the plausibility of the assumptions. These are considered in light of the present evaluation of the research methodology in this field.

Evidence. Evidence is often cited in support of the claim that rhyme and reading are causally connected (Goswami, 1990; 1991b; 1994; 1998; 1999a; 1999b; 2001). Although analysis in Appendix C of some of this evidence reveals little support for this claim, a number of teaching recommendations has been made (e.g. Goswami, 1999a, pp. 232–233): pre-reading instruction should focus on developing children's rhyme awareness skills; beginning reading instruction should help children to link rhyme (in speech) with rime units (in print), and to make use of rime units to read by analogy. How justified are these teaching recommendations?

First, will pre-school instruction in rhyme awareness lead to better reading progress? Examination of the research evidence and the quality of the methodology employed appears neither to support adequately the notion that rhyme awareness is correlated with reading ability, nor the claim that rhyme awareness training will lead to improved reading progress (see Claim 1 and 2). Furthermore, evidence suggests that the inclusion of rhyme awareness training, later on, as a part of a beginning reading programme is likely to have no effect whatsoever on children's reading ability. It may even hinder progress by necessitating a much slower pace of letter-sound instruction. Both correlational and training-study evidence, involving beginning readers, suggests that the more letter-sound teaching children receive at this stage (undiluted by rhyme awareness instruction), the better their reading progress (Solity, 1999; Stuart, 1999; Watson, 1998).

Second, there is no reliable evidence to date that teaching children how to link spoken rhyme segments with printed rime units, or how to use a rime analogy strategy will speed early reading progress over other forms of instruction (see Claim 2; also Table 7, type 3 studies).

Logical development sequence. The first part of the causal connections theory assumes that children's awareness of phonological units develops from large (rhyme) to small (phoneme) segments. Thus, it is recommended that instruction focus first on rhyme and rime units, followed by initial phonemes and onsets, a rime-analogy

reading strategy, and finally, a letter-sound reading strategy where rime units are segmented. How warranted is this advocated sequence of instruction?

This large-to-small unit theory of development appears to stem largely from results obtained from *implicit* rhyme tasks (such as the rhyme oddity measure) where an unconscious kind of global awareness of rhyme is needed, and at the same time, upon results obtained from *explicit* phoneme measures where phoneme units must be manipulated, or specifically identified (for discussion, Duncan et al, 1997, 2000; Seymour et al, 1999). Although a common view is that very young children are unable to perform any kind of phoneme task, there is evidence that this is not necessarily the case (Fox & Routh, 1975; Thomas & Senechal, 1998). Chaney (1992) found that among a sample of 3-year-olds, the percentage of children who met the passing criteria was 93%, 91%, and 88% on three phoneme tasks, all of which were implicit tasks, versus only 35% on an implicit, rhyme production task and only 26% on an explicit, rhyme identification task.

In fact, it appears that children must possess some letter-sound knowledge in order to perform phoneme or rhyme tasks that require an explicit awareness of these units. Since knowledge of a word's spelling assists performance on such tasks, the *explicit awareness* of these units refers more accurately to the ability to identify or manipulate, not phonemes or rhyme, but letter-sounds and rime units. Importantly, with the introduction of a phonemically-based alphabet once school starts, the explicit awareness of phonemes develops far more rapidly among beginning readers than does the explicit awareness of rhyme units. This is simply due to the nature of the English writing system. In a simple cvc word there are, not two symbols (one for the onset and one for the rime), but three letter symbols each separated in space, each symbol representing, and illuminating, a phoneme unit. Onsets and rime units, however, are not differentiated or *made explicit* by our writing system in this way.

Indeed, the *explicit awareness* of phonemes and rhyme units was measured among beginning readers in two experimental studies resulting in identical findings, in spite of differences between studies in the instructional methods to which children had been exposed (Duncan et al, 1997; Goswami & East, 2000). In both studies, the use of measures requiring *explicit awareness* revealed that children were far better at identifying initial and final phonemes in words than at identifying rhyme units. In fact, in the latter study, training designed to enhance the explicit awareness of rhyme units did not alter this picture. Although rhyme unit identification improved after training, children were still better at identifying the shared phonemes in pairs of spoken words, and when required to match their responses to written words, they were still better at underlining, not rime units, but the initial, medial and final letters in words. These children had spent just four months in school, having received instruction predominantly at the whole-word level, followed by five hours of rhyme and analogy training, and yet their brief acquaintance with our alphabetic writing system appeared to be sufficient to result in the superior availability of the phoneme.

Furthermore, Duncan and others (1997) found that children whose *implicit* awareness of rhyme was excellent were unable to identify rhyme units on a task requiring *explicit* awareness of rhyme. In short, this evidence tends to erode the argument for large-to-small unit sequencing of instruction, because as Seymour and others (1999) point out, the theory confuses the implicit and explicit awareness of phonological units; it assumes that the awareness of rhyme and the explicit identification of rhyme

units are the same, that the awareness of phonemes and the explicit identification of phonemes represent the same or a similar skill.

Even if there were no confusion over these different levels of linguistic awareness, it may be dangerous to design instruction according to a development view, according to what skills appear to be easier, when instruction targeted at developing the skills that are supposedly more difficult may be more effective. The developmental view, for example, assumes that initial phoneme detection is easier than final and medial phoneme detection and therefore, the suggestion made is that letter-sound teaching should begin with onsets. Yet, there is evidence that beginners are able to learn letter-sound correspondences at the same pace whether teaching is focused on phonemes in the onset position only, or on phonemes in all positions of a word (Watson, 1998, Experiment 2). And, nine months after training, compared to the latter type of instruction, onset-focused instruction results in significantly poorer letter-sound knowledge, *as poor as that of children who receive no letter-sound instruction of any kind*. Thus, the teaching of letter-sound correspondences through a focus on onsets alone may have no more effect than if this teaching did not occur at all.

Plausible assumptions. The rhyme and reading causal connection may owe part of its influence to the fact that it is based upon plausible assumptions. For example, it is assumed that because the spelling patterns of English are often irregular, beginning reading instruction that focuses on the more regular segments of words (the rime units) will facilitate reading progress over phoneme level teaching on its own. It is assumed, too, that phoneme awareness is important, even necessary, in order to learn how to read. But, to what extent do these assumptions bear analysis?

Although instruction in how to use a rime analogy strategy may help to highlight regularities in English spelling, it necessarily fails as a useful form of early reading instruction. The strategy cannot, in fact, be considered a beginning reading strategy because some letter-sound decoding skill and a considerable sight vocabulary are needed *first,* in order to use it. Additionally, the strategy has poor generative potential. To illustrate, in the UK, learning of the first fifteen rime units taught in the *National Literacy Strategy* along with five letter-sounds (i.e. 20 pieces of information) enables a child to read thirty-six words. Learning only half as many pieces of information at the phoneme level, however, (the same five letter-sounds plus five more) is sufficient to enable a child to read more than five times the number of words (183 words).

Finally, the fact that phoneme awareness and reading are highly correlated likely contributes to the plausible assumption that children must be able to detect phonemes in words in order to learn how to read. Yet, the bulk of current evidence suggests that while phonological training may improve phonological skills, it rarely has any effect on reading ability; in contrast, phonological training *combined with letter-sound teaching* can produce some positive effects. Bryant and others (1983), for example, found that children trained in both phonological and letter-sound skills were the only training group among three, to show significantly better progress in subsequent reading and spelling skills.

Interestingly though, the effects of combined phonological and letter-sound teaching are modest compared to when beginning reading instruction is concentrated on letter-sound teaching alone (see Claim 2 section). In fact, when one examines just what eye movement and brain processes occur during reading, this finding is entirely

consistent. When faced with the task of reading a cvc word, only three skills are involved: letter shape recognition, the left to right, letter-to-sound translation of each letter in turn, and the blending together of the three letter-sounds to pronounce the word. Phoneme segmentation skills are not required because the white spaces surrounding the three letters indicate where each of the three letter-to-sound translations should occur. Being able to detect phonemes in words is not required either; and blending of the three letter-sounds is facilitated by the presence of the letters, which function as visual memory prompts. Indeed, much evidence suggests that the awareness of phonemes only develops as a direct result of letter-sound instruction.

Moreover, although beyond the scope of this investigation, within the 61 studies evaluated here, there is at least some evidence suggesting that it is at later stages, beyond the beginning-to-read stage, that the explicit awareness of both phoneme and rhyme begin to develop and become more closely associated with reading age.

Acknowledgements

I would like to thank Chloë McComb, a graduate student currently working in the Department of Neuroscience, at the University of Calgary for her assistance in independently rating a sample of the studies evaluated here. I am grateful, as well, to all my reviewers for their helpful comments on an earlier version of this article.

References

Ackerman, P.T. & Dykman, R.A. (1993). Phonological processes, confrontational naming, and immediate memory in dyslexia. *Journal of Learning Disabilities*, 26, 597–609.

Adams, M. J. (1994). *Beginning to read: thinking and learning about print*. Cambridge, Mass: MIT Press.

Baker, L., Fernandez-Fein, S., Scher, D., & Williams, H. (1998). Home experiences related to the development of word recognition. In J.L. Metsala and L.C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 263–287). Hillsdale, NJ: Erlbaum.

Baron, J. (1977). Mechanisms for pronouncing printed words: use and acquisition. In D. LaBerge and S.J. Samuels (Eds.), *Basic processes in reading: perception and comprehension* (pp. 175–216). Hillsdale, NJ: Lawrence Erlbaum Associates.

Barron, R., Golden, J., Seldon, D., Tait, C., Marmurek, H., & Haines, L. (1992). Teaching prereading skills with a talking computer. *Reading and Writing: An Interdisciplinary Journal*, 4, 179–204.

Bowey, J. (1994). Phonological sensitivity in novice readers and nonreaders. *Journal of Experimental Child Psychology*, 58, 134–159.

Bowey, J. & Francis, J. (1991). Phonological analysis as a function of age and exposure to reading instruction. *Applied Psycholinguistics*, 12, 91–120.

Bradley, L. (1980). *Assessing reading difficulties*. London: Macmillan Educational.

Bradley, L. & Bryant, P. (1978). Difficulties in auditory organisation as a possible cause of reading backwardness. *Nature*, 271, 746–747.

Bradley, L. & Bryant, P. (1983). Categorizing sounds and learning to read. *Nature*, 301, 419–421.

Bradley, L. & Bryant, P. (1985). *Rhyme and reason in reading and spelling*. Michigan: University of Michigan Press.

Bruck, M. & Treiman, R. (1992). Learning to pronounce words: the limitations of analogies. *Reading Research Quarterly*, 27, 374–388.

Bryant, P. (1998). Sensitivity to onset and rhyme does predict young children's reading: A comment on Muter, Hulme, Snowling, and Taylor (1997). *Journal of Experimental Child Psychology*, 71, 29–37.

Bryant, P., Bradley, L., Maclean, M., & Crossland, J. (1989). Nursery rhymes, phonological skills and reading. *Journal of Child Language*, 16, 407–428.

Bryant, P., & Goswami, U. (1986). Strengths and weaknesses of the reading level design – comment on Backman, Mamen, and Ferguson. *Psychological Bulletin*, 100, 101–103.

Bryant, P., Maclean, L., Bradley, L., & Crossland, J. (1990a). Rhyme and alliteration, phoneme detection, and learning to read. *Developmental Psychology*, 26, 429–438.

Bryant, P., Maclean, L. & Bradley, L. (1990b). Rhyme, language, and children's reading. *Applied Psycholinguistics*, 11, 237–252.

Burt, L., Holm, A. & Dodd, B. (1999). Phonological skills of 4-year-old British children: an assessment and developmental data. *International Journal of Language and Communication Disorders*, 34, 311–335.

Calfee, R.C., Lindamood, P. & Lindamood, C. (1973). Acoustic-phonetic skills and reading: Kindergarten through twelfth grade. *Journal of Educational Psychology*, 64, 293–298.

Campbell, D.T. & Stanley, J.C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Carduso-Martins, C. (1994). Rhyme perception: Global or analytical? *Journal of Experimental Child Psychology*, 57, 26–41.

Carillo, M. (1994). Development of phonological awareness and reading acquisition. *Reading and Writing: An Interdisciplinary Journal*, 6, 279–298.

Chaney, C. (1992). Language development, metalinguistic skills and print awareness in 3-year-old children. *Applied Psycholinguistics*, 13, 485–514.

Chaney, C. (1994). Language development, metalinguistic awareness and emergent literacy skills of 5-year-old children in relation to social class. *Applied Psycholinguistics*, 15, 371–374.

Cronin, V. & Carver, P. (1998). Phonological sensitivity, rapid naming, and beginning reading. *Applied Psycholinguistics*, 19, 447–461.

Cunningham, P.M. (1976). Investigating a synthesised theory of mediated word identification. *Reading Research Quarterly*, 11, 127–143.

Deavers, R., Solity, J.E., & Kerfoot, S. (2000). The effect of instruction on early nonword reading strategies. *Journal of Research in Reading*, 23, 267–286.

De Jong, P. F. & Van Der Leij, A. (1999). Specific contribution of phonological abilities to early reading acquisition: Results from a Dutch latent variable longitudinal study. *Journal of Educational Psychology*, 91, 450–476.

DfEE (1998). *The National Literacy Strategy: framework for teaching*. London: Department of Education and Employment.

Duncan, L. & Johnston, R. (1999). How does phonological awareness relate to nonword reading skill amongst poor readers? *Reading and Writing: An Interdisciplinary Journal*, 11, 405–439.

Duncan, L. & Seymour, P. (2000a). Phonemes and rhyme in the development of reading and metaphonology: The Dundee longitudinal study. In N. Badian (Ed.), *Prediction and prevention of reading failure* (pp. 275–297). Parkton, Maryland: York Press.

Duncan, L. & Seymour, P. (2000b). Socio-economic differences in foundation-level literacy. *British Journal of Psychology*, 91, 145–166.

Duncan, L., Seymour, P., & Hill, S. (1997). How important are rhyme and analogy in beginning reading? *Cognition*, 63, 171–208.

Duncan, L., Seymour, P., & Hill, S. (2000). A small-to-large unit progression in metaphonological awareness and reading? *The Quarterly Journal of Experimental Psychology*, 53A, 1081–1104.

Ellis, N. & Large, B. (1987). The development of reading: as you seek so shall you find. *British Journal of Psycholinguistics*, 78, 1–28.

Fayne, H. & Bryant, D. (1981). Relative effects of various word synthesis strategies on the phonics achievement of learning disabled youngsters. *Journal of Educational Psychology*, 73, 616–623.

Fernandez-Fein, S. & Baker, L. (1997). Rhyme and alliteration sensitivity and relevant experiences in preschoolers from diverse backgrounds. *Journal of Literacy Research*, 29, 433–459.

Fox, B. & Routh, D. (1975). Analyzing spoken language into words, syllables, and phonemes: A developmental study. *Journal of Psycholinguistic Research*, 3, 331–342.

Gaskins, R.W., Gaskins, I.W., Anderson, R.C. & Schommer, M. (1995). The reciprocal relationship between research and development: An example involving a decoding strand for poor readers. *Journal of Reading Behaviour*, 27, 337–377.

Gathercole, S. & Willis, C. (1991). Differentiating phonological memory and awareness of rhyme: reading and vocabulary development in children. *British Journal of Psychology*, 82, 387–406.

Goswami, U. (1988). Orthographic analogies and reading development. *The Quarterly Journal of Experimental Psychology*, 40A, 239–268.

Goswami, U. (1990). A special link between rhyming skill and the use of orthographic analogies by beginning readers. *Journal of Child Psychology and Psychiatry*, 31, 301–311.

Goswami, U. (1991b). Learning about spelling sequences: the role of onsets and rimes in analogies in reading. *Child Development*, 62, 1110–1123.
Goswami, U. (1991b). *Phonological skills and learning to read*. Hove: Lawrence Erlbaum.
Goswami, U. (1994). Reading by analogy: theoretical and practical perspectives. In C. Hulme & M. Snowling (Eds.), *Reading development and dyslexia*. Whurr Publishers.
Goswami, U. (1995). Phonological development and reading by analogy: What is analogy, and what is it not? *Journal of Research in Reading*, 18, 139–145.
Goswami, U. (1998). The role of analogies in the development of word recognition. In J.L. Metsala & L.C. Ehri (Eds.) *Word recognition and beginning literacy*. Mahwah, NJ: Lawrence Erlbaum Associates.
Goswami, U. (1999a). Causal connections in beginning reading: the importance of rhyme. *Journal of Research in Reading*, 22, 217–240.
Goswami, U. (1999b). Phonological development and reading by analogy: epilinguistic and metalinguistic issues. In J. Oakhill & R. Beard (Eds.), *Reading development and the teaching of reading: A psychological perspective*. Malden, Mass: Blackwell Publishers.
Goswami, U. (2001). Rhymes are important: a comment on Savage. *Journal of Research in Reading*, 24, 19–29.
Goswami, U. & Bryant, P. (1990). *Phonological Skills and Learning to Read*. Hove: Lawrence Erlbaum.
Goswami, U. & East, M. (2000). Rhyme and analogy in beginning reading: conceptual and methodological issues. *Applied Psycholinguistics*, 21, 63–93.
Greaney, K.T. & Tunmer, W.E. (1996). Onset-rime sensitivity and orthographic analogies in normal and poor readers. *Applied Psycholinguistics*, 17, 15–40.
Greaney, K.T. & Tunmer, W.E. & Chapman, J.W. (1997). Effects of rime-based orthographic analogy training on the word recognition skills of children with reading disability. *Journal of Educational Psychology*, 89, 645–651.
Hansen, J. & Bowey, J. (1994). Phonological analysis skills, verbal working memory, and reading ability in second-grade children. *Child Development*, 65, 938–950.
Haskell, D., Foorman, B., & Swank, P. (1992). Effects of three orthographic-phonological units on first-grade reading. *Remedial and Special Education*, 13, 40–49.
Hatcher, P. & Hulme, C. (1999). Phonemes, rhymes, and intelligence as predictors of children's responsiveness to remedial reading instruction: Evidence from a longitudinal intervention study. *Journal of Experimental Child Psychology*, 72, 130–153.
Hoien, T., Lundberg, I., Stanovich, K. & Bjaalid, I. (1995). Components of phonological awareness. *Reading and Writing: An Interdisciplinary Journal*, 7, 171–188.
Hurford, D., Schauf, J., Bunce, L., Blaich, T. & Moore, K. (1994). Early identification of children at-risk for reading disabilities. *Journal of Reading Disabilities*, 27, 371–382.
Johnston, R. & Rugg, M. (1989). Rhyme judgement ability in good and poor readers. *Language and Education*, 3, 223–233.
Johnston, R., Anderson, M., & Holligan, C. (1996). Knowledge of the alphabet and explicit awareness of phonemes in pre-readers: The nature of the relationship. *Reading and Writing*, 8, 217–234.
Joseph, L.M. (2000). Developing first graders' phonemic awareness, word identification and spelling: A comparison of two contemporary phonic instructional approaches. *Reading Research and Instruction*, 39, 160–169.
Kirtley, C., Bryant, P., Maclean, M., & Bradley, L. (1989). Rhyme, rime, and the onset of reading. *Journal of Experimental Child Psychology*, 48, 224–245.
Leslie, L. & Allen, L. (1999). Factors that predict success in an early literacy intervention project. *Reading Research Quarterly*, *34*, 404- 424.
Levy, B., Bourassa, D., & Horn, C. (1999). Fast and slow namers: Benefits of segmentation and whole word training. *Journal of Experimental Child Psychology*, 73, 115–138.
Lundberg, I., Olaffson, & Wall, S. (1980). Reading and spelling skills in the first school years predicted from phonemic awareness kills in kindergarten. *Scandinavian Journal of Psychology*, 21, 159–173.
Lundberg, I., Frost, J., & Petersen, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly*, 23, 263–284.
Maclean, M., Bryant, P. E., & Bradley, L. (1987). Rhymes, nursery rhymes and reading in early childhood. *Merrill-Palmer Quarterly*, 33, 255–282.
Majsterek, D.J. & Ellenwood, A.E. (1995). Phonological awareness and beginning reading: Evaluation of a school-based screening procedure. *Journal of Learning Disabilities*, 28, 449–456.
McDougall, S., Hulme, C. Ellis, A., & Monk, A. (1994). Learning to read: The role of short-term memory and phonological skills. *Journal of Experimental Child Psychology*, 58, 112–133.

McGuinness, D., McGuinness, C., & Donohue, J. (1995). Phonological training and the alphabet principle: Evidence for reciprocal causality. *Reading Research Quarterly*, 30, 830–852.
Muter, V., Hulme, C., Snowling, M., & Taylor, S. (1997). Segmentation, not rhyming, predicts early progress in learning to read. *Journal of Experimental Child Psychology*, 65, 370–396.
Muter, V., Hulme, C., Snowling, M., & Taylor, S. (1998). Segmentation, not rhyming, predicts early progress in learning to read. *Journal of Experimental Child Psychology*, 71, 3–27.
Muter, V. & Snowling, M. (1998). Concurrent and longitudinal predictor of reading: The role of metalinguistic and short-term memory skills. *Reading Research Quarterly*, 33, 320–337.
Naslund, J. & Schneider, W. (1996). Kindergarten letter knowledge, phonological skills and memory processes: relative effects on early literacy. *Journal of Experimental Child Psychology*, 62, 30–59.
Nation, K & Hulme, C. (1997). Phonemic segmentation, not onset-rime segmentation, predicts early reading and spelling skills. *Reading Research Quarterly*, 32, 154–167.
Oakhill, J. & Kyle, F. (2000). The relation between phonological awareness and working memory. *Journal of Experimental Child Psychology*, 75, 152–164.
O'Connor, R.E., Slocum, T.A. & Jenkins, J.R. (1995). Transfer among phonological tasks in kindergarten: Essential instructional content. *Journal of Educational Psychology*, 87, 202–217.
O'Shaughnessy, T.E. & Swanson, H.L. (2000). A comparison of two reading interventions for children with reading disabilities. *Journal of Learning Disabilities*, 33, 257–277.
Peterson, M.L. & Haines, L. (1992). Orthographic analogy training with kindergarten children: Effects on analogy use, phonemic segmentation, and letter-sound knowledge. *Journal of Reading Behaviour*, 24, 109–126.
Qi, S. & O'Connor, R. (2000). Comparison of phonological training procedures in kindergarten classrooms. *Journal of Educational Research*, 93, 226–233.
Rohl, M. and Pratt, C. (1996). Phonological awareness, verbal working memory and the acquisition of literacy. *Reading and Writing*, 8, 327–360.
Santa, C. & Hoien, T. (1999). As assessment of early steps: A program for early intervention of reading problems. *Reading Research Quarterly*, 34, 54–73.
Scarborough, H.S.L. (1990). Very early language deficits in dyslexic children. *Child Development*, 61, 1728–1743.
Schult, R. (1999). *Investigating the social world* (2nd ed.). Pineforge Press.
Seymour, P., Duncan, L. & Bolik, F. (1999). Rhymes and phonemes in the common unit task: replications and implications for beginning reading. *Journal of Research in Reading*, 22, 113–130.
Snowling, M., Hulme, C., Smith, A. & Thomas, J. (1994). The effect of phonetic similarity and list length on children's sound categorisation performance. *Journal of Experimental Child Psychology*, 58, 16–180.
Solity, J., Deavers, R., Kerfoot, S., Crane, G. & Cannon, K. (1999). Raising literacy attainments in the early years: The impact of instructional psychology. *Educational Psychology*, 19, 373–398.
Spatz, C. and Johnston, J. (1984). *Basic statistics 3rd edition tales of distribution*. Monterey: Brooks/Cole.
Stahl, S. and Murray, B. (1994). Defining phonological awareness and its relationship to early reading. *Journal of Educational Psychology*, 86, 221–234.
Stainthorp, R. and Hughes, D. (1998). Phonological sensitivity and reading: evidence from precocious readers. *Journal of Research in Reading*, 21, 53–68.
Stanovich, K., Cunningham, A. and Cramer, B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Child Psychology*, 38, 175–190.
Stuart, M. (1995). Prediction and qualitative assessment of five-and six-year-old children's reading: a longitudinal study. *British Journal of Educational* Psychology, 65 (part 3), 287–296.
Stuart, M. (1999). Getting ready for reading: early phoneme awareness and phonics teaching improves reading and spelling in inner-city second language learners. *British Journal of Educational Psychology*, 69, 587–605.
Sullivan, H.J., Okada, M, & Niedermeyer, F.C. (1971). Learning and transfer under two methods of word-attack instruction. *American Educational Research Journal*, 8, 227–239.
Tabachnick , B.G. and Fidell, L.S. (1996). *Using multivariate statistics*. Mass: Allyn Bacon.
Tabachnick , B.G. and Fidell, L.S. (2001). *Using multivariate statistics*. Mass: Allyn Bacon
Thomas, E.M. & Senechal, M. (1998). Articulation and phoneme awareness of 3-year-old children. *Applied Psycholinguistics*, 19, 363–391.
Troia, G. (1999). Phonological awareness intervention research: A critical review of the experimental methodology. *Reading Research Quarterly*, 34, 28–52.

Vadasy, P., Jenkins, J. & Pool, K. (2000). Effects of tutoring in phonological and early reading skills on students at risk for reading disabilities. *Journal of Learning Disabilities*, 33, 579–590.

Wagner, R., Torgesen, J., Rashotte, C., Hecht, S., Barker, T., Burgess, S. Donahue, J. & Garon, T. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: a 5-year longitudinal study. *Developmental Psychology*, 33, 468–479.

Walton, P.D. (1995). Rhyming ability, phonemic identity, letter-sound knowledge and the use of orthographic analogy by prereaders. *Journal of Educational Psychology*, 87, 587–597.

Walton, P.D. & Walton, L.M. (2001). Beginning reading by teaching in rime analogy: Effects on phonoligcal skills, letter-sound knowledge, working memory, and word reading strategies. *Scientific Studies of Reading*, (in press).

Walton, P.D., Walton, L.M., & Felton, K. (2000). Teaching rime analogy or letter recoding reading strategies to prereaders: effects on prereading skills and word reading. *Journal of Educational Psychology*, (in press).

Walton, P.D., Thornloe, S., Bowden, M. E., Angus, M. (2001). Evaluation of a rime-based reading program with Shuswap and Heiltsuk first nations prereaders. *Reading and Writing*, (in press).

Watson, J.E. (1999). *An investigation of the effects of phonics teaching on children's progress in reading and spelling*. Doctoral dissertation, University of St Andrews, St. Andrews, Scotland.

Webster, P.E. & Plante, A.S. (1992). Phonologically impaired preschoolers: Rhyme with an eye towards reading. *Perceptual and Motor Skills*, 75, 1195–1204.

Webster, P.E. & Plante, A.S. (1995). Productive phonology and phonological awareness in preschool children. *Applied psycholinguistics*, 16, 43–57.

Wilson, J. & Frederickson, N. (1995). Phonological awareness training: An evaluation. *Educational and Child Psychology*, 12, 68–79.

Wise, B. & Olson, R. (1990). Subsyllabic units in computerized reading instruction: Onset-rime vs. postvowel segmentation. *Journal of Experimental Child Psychology*, 49, 1–19.

Wise, B., Ring, J., & Oolson, R.K. (1999). Training phonological awareness with and without explicit attention to articulation. *Journal of Experimental Child Psychology*, 72, 271–304.

Wise, B., Ring, J., & Olson, R.K. (2000). Individual differences in gains from computer-assisted remedial reading. *Journal of Experimental Child Psychology*, 77, 197–235.

Wilkinson, L. and the Task Force on Statistical Inference (1999). Statistical methods in psychology journals: guidelines and explanations. *American Psychologist*, 54.

Wood, C. & Terrell, C. (1998). Pre-school phonological awareness and subsequent literacy development. *Educational Psychology*, 18, 253–274.

Yopp, H. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly*, *23*, 159–176.

Address for correspondence: Dr Bonnie Macmillan, University of Birmingham, School of Psychology, Egbaston, Birmingham B15 2TT, UK. E-Mail: *bmacmillan @rivu.demon.co.uk*

Appendix A. A Selection of Training Studies Excluded on the Basis of Various Criteria

Exclusion Criteria	Study
No control group, unseen control, or control group did not receive training	Gaskins et al (1995); Goswami (1988); (1990); (1991a); Goswami & East (2000); Leslie & Allen (1999); Peterson & Haines (1992); Vadasy et al (2000); Walton et al (2001) Exp. 2; Wilson & Frederickson (1995)
No general reading measure scores as the dependent variable	Goswami (1988)*; (1990)*; (1991a)*; Goswami & East (2000)*
Inadequately stringent comparison: (rhyme/rime training with letters vs. other training without letters	Bradley and Bryant (1983) (rhyme + plastic letters group vs. other); de Jong et al (2000); Hatcher, Hulme & Ellis (1994); Walton et al (2001) Exp. 1; Wise et al (2000)
Inadequate or biased post-treatment reading measures; in-house reading measures only	Bruck & Treiman (1992); Haskell, Foorman & Swank (1992); Fayne & Bryant (1981); Levy et al (1999); O'Connor et al (1995); Sullivan et al (1971); Walton (1995); Walton, Walton, & Felton (2000); Walton & Walton (2001); Walton et al (2000) Exp. 1 and 2; Wise et al (1990); Wise & Olson (195); Wise et al (1999); Wise et al (2000)

* This study was primarily aimed at testing a developmental reading model, and was not directly concerned with the effects of analogy training on reading in general.

Appendix B. Study type and sample composition among 32 correlational studies evaluated

	Correlational studies involving concurrent measurement **(purpose to establish association)**	**Predictive studies involving longitudinal measurement** **(purpose to establish time order association)**
Non-readers/ pre-school subjects	Gathercole & Willis, 1991; Johnston et al, 1996; Rohl & Pratt, 1996	Bryant et al, 1990a; De Jong & Van der Leij, 1999; Lundberg et al 1980; Muter & Snowling , 1998; Naslund & Schneider, 1996; Stanovich et al, 1984; Stuart, 1999; Wagner et al, 1997; Wood & Terrell, 1998
Beginning readers/ older readers	Bowey & Francis, 1991; Carduso-Martins, 1994 (p. 34 only); Hansen & Bowey, 1994; Hoien et al, 1995 (study 2); Kirtley et al, 1989 (Exp.2); McDougall, 1994 (analyses 1, 2, 4); McDougall, 1994 (analysis 3)	Bradley & Bryant, 1983; Duncan et al 1997; Hatcher & Hulme, 1999; Muter et al, 1997
Different reading ability groups	Bradley & Bryant, 1978; Carduso-Martins, 1994 (Kindergarten only, p 32–33); Carillo, 1994; Cronin & Carver, 1998; Duncan & Johnston, 1999; Johnston & Rugg, 1989; Stainthorp & Hughes, 1980	Ellis & Large, 1987; Nation & Hulme, 1997

Appendix C. Analysis of 19 studies* cited by Goswami (1999, p. 219) in support of 'the early availability of rhyme and its causal links with reading'

Type of evidence	Topic	Study
A. Availability of rhyme: Evidence establishing the 'early availability of rhyme'	Very young children demonstrate an awareness of rhyme, rhyme combined with something else, or knowledge of nursery rhymes	2, 4, 5, 6, 7, 8, 9, 10, 11, 14, 15, 17, 18, 19
B. Causation: Evidence establishing that there are 'causal links' between rhyme and reading; evidence from both correlational and training studies is required	Evidence excluded because – a) relationship between 2 variables, but not rhyme, and not reading b) relationship between something else (not rhyme) or between rhyme contaminated with something else and reading c) not relevant (a review only, or reading not the dependent variable) d) rhyme awareness is *not* related to reading ability e) rhyme or rime-related training does *not* improve reading ability	7 8 11 1 2a 5 9a 13 15 16 3 14 9b 2b 17
	Evidence remaining – 1) Correlational evidence: a) rhyme and reading, measured concurrently are related; b) rhyme and reading, measured longitudinally, are related. 2) Training study evidence: a) rime, not rhyme-related training produces positive effects on reading ability.	4 6 10 12

SUMMARY

A. Early availability of rhyme: 15 out of 19 studies provide some evidence.

B. Causal connection between rhyme and reading: 4 out of 19 studies provide some partial evidence (#4, 6, 10, 12). However, the three correlational studies (#4, 6 and 10) fail to meet a number of important research methodology criteria (see Table 2) seriously undermining the validity of findings. For example, none of these studies properly controlled for guessing, had normally distributed independent variable scores, had adequate subject to variable ratios, or controlled for the effects of alphabet knowledge. The only training study (#12), compares rime (not rhyme) training with training in the use of context cues, or in effect, compares some alphabetic instruction with very little or none.

* The 19 studies (numbered for this analysis):
1) Bradley & Bryant (1978) 2a) Bradley & Bryant, Exp 1 (1983) 2b) Bradley & Bryant, Exp 2 (1983) 3) Baker et al (1998) 4) Bowey & Francis (1991) 5) Bryant et al (1989) 6) Bryant et al (1990) 7) Chaney (1992) 8) Chaney (1994) 9a) Cronin & Carver, 1st analysis (1998) 9b) Cronin & Carver, 2nd analysis (1998) 10) Ellis & Large (1987) 11) Fernandez-Fein & Baker (1997) 12) Greaney et al (1997) 13) Hansen & Bowey (1994) 14) Maclean et al (1987) 15) Scarborough (1990) 16) Stahl & Murray (1994) 17) Walton (1995) 18) Webster & Plante (1992) 19) Webster & Plante (1995)