Learning to Spell: Variability, Choice, and Change in Children’s Strategy Use

Bethany Rittle-Johnson and Robert S. Siegler

We examined whether the overlapping waves model, originally developed to account for strategy choices in arithmetic, could also account for strategy choices in spelling. The contrast was of particular interest because arithmetic is an algorithmic domain (a domain that includes strategies that always yield correct answers if executed properly), whereas spelling is not. Thirty first-grade students spelled words under 2 conditions, and 23 of these students were retested in second grade. Trial-by-trial analysis of strategy use was used to identify which strategies first and second graders used, how adaptively they chose among them, how effective the strategies were, and what changes occurred from first to second grade along each dimension. The model proved useful for understanding the development of spelling, despite the fact that explicit use of backup strategies had a minimal impact on accuracy. Implications for understanding adaptive strategy choices in algorithmic and nonalgorithmic domains are discussed.

INTRODUCTION

Cognitive development is often depicted as a sequence of distinct levels of thought. At early ages, thought is governed by a certain stage, strategy, theory, rule, or form of representation. At later ages, children progress to higher and higher levels. These theories highlight differences in children’s thinking at various ages, but they offer limited insight into the process of change.

The overlapping waves model (Siegler, 1996) is based on a different set of premises. It starts with the view that on most tasks, children think in a variety of ways, not just one. The assumption that children’s thinking is highly variable leads to a focus on how children choose among the varied approaches. It also leads to a focus on gradual changes in the distributions of ways of thinking and in the efficiency with which various strategies are executed, as well as on the introduction of new approaches. Through trial-by-trial examination of children’s strategy use and how it changes with experience, this theoretical approach has yielded more precise descriptions of development in a number of domains than were heretofore available.

In the present study, we applied the overlapping waves approach to first and second graders’ spelling. Trial-by-trial analyses of overt behavior and immediately retrospective self-reports were used to identify what strategies first graders used to spell each word, how adaptively they chose among the strategies, how effective the strategies were, and how their strategy use changed after a year of spelling experience. The main goals were to provide a more precise description of first and second graders’ spelling strategies than currently available and to test whether the overlapping waves model is generalizable to a different class of domains than those to which it has been applied previously.

The Overlapping Waves Model

Within the overlapping waves model, abundant variability, adaptive choice, and gradual change are fundamental features of cognition at all points in development. First consider evidence regarding variability. In such diverse domains as arithmetic, map drawing, scientific reasoning, conservation, and motor development, infants, preschoolers, elementary school students, and college students have been found to use multiple strategies (Adolph, Eppler, & Gibson, 1993; Church & Goldin-Meadow, 1986; Feldman, 1980; Geary & Burlingham-Dubree, 1989; Kuhn, Garcia-Mila, Zohar, & Anderson, 1995; Siegler & Shrager, 1984). This variability is often present from the beginning of development within a domain. For example, French children in the first two weeks of learning single-digit multiplication use an average of three strategies to multiply (Lemaire & Siegler, 1995).

Knowing diverse strategies allows children to adapt their strategy use to task demands. For example, when solving arithmetic problems, children usually select faster and less effortful strategies on easier problems and slower and more effortful strategies on more difficult ones (Siegler, 1986). Such choices allow them to proceed quickly and accurately when that is possible, and more slowly but still accurately when the problems are too difficult to be solved quickly and accurately. Children also adjust their strategy use to task demands by flexibly combining several strategies to solve a problem. For example, elementary school children often used two or three memory strategies while studying a single set of words (Coyle & Bjorklund, 1997).

© 1999 by the Society for Research in Child Development, Inc.
All rights reserved. 0009-3920/99/7002-0006
Within the overlapping waves approach, change toward faster and more accurate performance is viewed as generally proceeding gradually. The changes can be produced in at least four ways: (1) introduction of new, more advanced, strategies, (2) increasing use of the more advanced strategies from among those that are already known, (3) increasingly effective execution of strategies, and (4) more adaptive choices among strategies. In one study in which all four potential sources of change were examined, large improvements in multiplication performance over the course of a school year resulted from increasing use of the fastest and most accurate strategy, increasingly effective execution of strategies, and increasingly adaptive choices among strategies (Lemaire & Siegler, 1995). Improvements did not stem from introduction of new strategies—the set of strategies that children used did not change.

Part of the reason for focusing on cognitive variability is its relation to cognitive change. Thinking and acting in varied ways gives children a chance to explore the task environment and to observe the consequences of alternative approaches. Consistent with this perspective, variability of pretest strategy use has been found to be positively related to amount of benefit from instruction on both mathematical equivalence and number conservation problems (Alibali & Goldin-Meadow, 1993; Perry, Church, & Goldin-Meadow, 1988; Siegler, 1995).

Algorithmic and Nonalgorithmic Domains

Past applications of the overlapping waves approach have focused on algorithmic domains, that is, domains that allow algorithmic strategies. An algorithmic strategy is a procedure that, if executed correctly, always yields the correct answer. For example, multidigit subtraction is algorithmic, because a child who correctly executes the subtraction procedure will always answer correctly.

Research in algorithmic domains has consistently shown that children choose adaptively among alternative strategies. Most of the demonstrations have involved the choice between retrieving an answer and using a backup strategy (an explicit strategy other than retrieval, such as adding by counting from one or by counting from the larger addend). When children face this choice, the harder the problem, the more likely they are to use a backup strategy. Such choices are sensitive to problem difficulty and adaptive in the sense that they allow quick and accurate performance when that is possible, and slower, but still accurate performance, when the problem is too difficult to be solved quickly and accurately. Such adaptive choices have been demonstrated in addition (Geary, 1990; Siegler & Shrager, 1984), subtraction (Siegler, 1987a), multiplication (Brownell & Carper, 1943; Cooney, Swanson, & Ladd, 1988), time telling (Siegler & McGilly, 1989), and mechanics problems in physics (Maloney & Siegler, 1993). They have been demonstrated among children with and without learning disabilities (Geary & Brown, 1991; Geary, Brown, & Samaranayake, 1991), among low income and middle income children (Kerkman & Siegler, 1993), among older and younger adults (Geary & Wiley, 1991; LeFevre, Sadesky, & Bisanz, 1996), and among Chinese, French, and American children (Geary, Fan, & Bow-Thomas, 1992; Lemaire & Siegler, 1995).

However, not all domains are algorithmic; in fact, most are not. In nonalgorithmic domains, no strategy guarantees success. Memorizing facts, generating scientific hypotheses, and spelling words in English without a dictionary are three examples of nonalgorithmic activities. Even a child who perfectly applied all of the orthographic regularities in the English language would misspell many words (Venezky, 1970).

The conceptual distinction between algorithmic and nonalgorithmic domains has not been made previously. Therefore, the few studies that have been done in nonalgorithmic domains, in particular memory development (Coyle & Bjorklund, 1997; McGilly & Siegler, 1990), were not set up to allow direct comparisons of performance in the two types of domains. For example, studies in algorithmic and nonalgorithmic domains have not used the same methods to examine the adaptiveness of strategy choices. This precluded direct comparison of the adaptiveness of choices in the two types of domains.

Such comparisons are particularly important because the uncertain effectiveness of backup strategies in nonalgorithmic domains could have large effects on the adaptiveness of choices between retrieval and backup strategies. The exact effect would depend on how adaptiveness is defined and on the mechanism for choosing between strategies. One definition of adaptive choice between retrieval and backup strategies is that choices follow the form: the less likely that retrieval will yield a correct answer on a problem, the more likely that backup strategies will be used on the problem. This view would follow from a strategy choice mechanism in which the probability of successful retrieval dictated the choice of strategy, regardless of the effectiveness of the backup strategies. Within this view, choices should be just as adaptive in nonalgorithmic domains as in algorithmic domains. Alternatively, choices between retrieval and backup strategies may be defined as adaptive only if they increase accuracy and speed beyond that which would be realized by consistent
reliance on retrieval alone or on backup strategies alone. This view would follow from a strategy choice mechanism that incorporates the speeds and accuracies of the backup strategies as well as of retrieval to select the strategies that are most likely to yield the correct solution in the shortest amount of time.

In algorithmic domains, the two senses of adaptiveness and the two mechanisms make similar predictions—the more difficult the problem, the more likely that backup strategies will be used. In nonalgorithmic domains, however, where backup strategies may not yield better outcomes than retrieval on difficult problems, the predictions that follow from the two views diverge. Within the first scenario, strategy choices would be equally sensitive to problem difficulty in nonalgorithmic as in algorithmic domains. Within the second scenario, strategy choices would be expected to be less adaptive in nonalgorithmic domains, to the extent that the backup strategies do not yield greater accuracy than retrieval. Thus, the pattern of strategy choices in a nonalgorithmic domain seemed likely to be informative regarding the mechanism by which strategy choices are made.

Strategy Choices in Spelling

Is the abundant variability, adaptive strategy choice, and gradual change found in algorithmic domains also present in nonalgorithmic domains such as spelling? Past models of spelling development suggest that it is not. Children have usually been depicted as progressing through a sequence of stages of strategy use and orthographic understanding (e.g., Ehri, 1992; Gentry, 1982; Henderson, 1980; Marsh, Friedman, Welch, & Desberg, 1980). Although differing in their particulars, all of the models propose that children first spell by matching the initial, or initial and final, sounds of a word to appropriate letters. In the next stage, children represent every sound they hear in a word on a one-sound/one-letter basis. Still later, in second or third grade, children begin to utilize orthographic rules (e.g., marking long vowels), analogies to other words, information about root words, and morphological endings, as well as sound patterns (Beers, 1980; Ehri, 1980; Evans & Smith, 1989; Gentry, 1982; Henderson, 1980; Turner & Quinn, 1986; Varnhagen, McCallum, & Burstow, 1997; Waters, Bruck, & Malus-Abramowitz, 1988; Zutell, 1980).

These models focus on the development of explicit problem solving strategies; they do not integrate the role of retrieval in learning to spell. Beginning spellers retrieve spellings of their names and other very high frequency words (Temple, Nathan, Temple, & Burris, 1993), and it seems likely that beginning spellers may use retrieval quite often. Thus, a complete model of spelling development would need to incorporate the role of retrieval in improved spelling ability.

We define retrieval as the rapid, automatic or close to automatic, activation of a spelling. This activation can rely on both word and subword (e.g., syllable, letter) information, including rapid phonological and morphological processing. It is not under conscious control and does not involve explicit application of rules (Plaut, McClelland, Seidenberg, & Patterson, 1996; Treiman, 1993; Treiman, Cassar, & Zukowski, 1994). In contrast, we define backup spelling strategies as explicit, controlled, step-by-step methods for constructing the sequence of letters. In general, they are slower than retrieval. For example, explicit sounding out typically involves the slow pronunciation of isolated phonemes and the step-by-step matching of phonemes to appropriate graphemes. Both empirical accounts and computational models support the separation of fluid, automatic processing from controlled, conscious use of problem solving strategies (Shallice, Glasspool, & Houghton, 1995; Shiffrin & Schneider, 1977). Of course, retrieval and backup strategies can be used in combination to generate the spelling of a word—a child may retrieve part of a spelling and sound out the rest—but the distinction between the two types of approaches captures an important difference in how spellings are produced.

Even if beginning spellers use multiple strategies, they may not choose adaptively among them. As noted above, in algorithmic domains, if children can reliably retrieve an answer, they usually do so; if the accuracy of the retrieved answer is uncertain, they generally use a slower backup strategy that is more likely to yield the correct answer. However, it is unclear if backup strategies in spelling, such as explicit sounding out, yield more accurate answers than only relying on retrieval. This uncertainty is especially great for beginning spellers, whose execution of backup strategies such as sounding out may be inaccurate even for words with regular sound-symbol correspondences. On difficult words, children are faced with choosing between retrieving a best guess or using backup strategies that also ultimately produce best guesses. If execution of backup strategies is so inaccurate that their use has little effect on spelling accuracy, children may not use them very much, because they require more time and effort to execute than does retrieval. Alternatively, beginning spellers may use backup strategies, but they may not use them disproportionately on the more difficult words. Yet a third possibility is that they use backup strategies disproportionately on difficult words, but without use of the backup strategies leading to increased accuracy.
It also is unclear whether spelling acquisition follows a path of gradual change. The stage models depict development of spelling as progressing through distinct stages in which introduction of a more advanced strategy results in the child moving to the next stage. These models imply that change occurs over a relatively short period of time, during which children discover a new strategy and use it whenever it is applicable. In algorithmic domains, trial-by-trial analysis has revealed that such stage models are inaccurate and that gradual changes are the rule (Siegler, 1996). When a new strategy is discovered, it typically is used only occasionally at first. Many changes occur not through introduction of new strategies but instead through increasingly efficient execution of existing strategies, increasing use of the more accurate strategies, and more adaptive choices among strategies. A more detailed examination of spelling strategies was expected to reveal similar gradual changes along these dimensions.

Past research on spelling development has not provided evidence needed to evaluate whether variability, adaptive choice, and gradual change are characteristic of young children's spelling strategies. Almost all of this past research is based on post hoc examination of spelling errors, with the notable exception of recent research by Steffler, Varnhagen, Friesen, and Treiman (1998). For example, use of sounding out often is inferred from spelling errors that contain plausible, but incorrect, letters. However, the errors could have been generated by an analogy to a particular word or by inaccurate retrieval of a previous misspelling. No strategy infersences whatsoever can be made from correct spellings, because any strategy could have yielded them. In many studies, these correct spellings account for 50% or more of young children's performance (Gentry, 1982). Thus, post hoc analyses do not allow accurate assessment of which strategies children use, how often and adaptively they use them, how accurately they use them, and how each of these aspects develops.

The Present Study

In this study, we assessed beginning spellers' strategies through two means: observations of ongoing behavior while spelling words and immediately retrospective verbal reports of strategy use. Because overt behavior is particularly clear evidence of strategy use when it is present, but does not accompany all instances of strategy use, combining the two measures provides a more sensitive and accurate measure of strategy use than simply relying on either alone. Past studies (McGilly & Siegler, 1990; Siegler, 1987b; Steffler et al, 1998) have shown that such assessments of strategy use are accurate and do not alter strategy use. The present study includes a new type of analysis that provides further evidence for the validity of this strategy assessment method.

In addition to assessing children's performance under conditions in which they were allowed to choose which strategies to use, we also assessed the same children's performance when use of explicit backup strategies was prohibited. The goal was for children to retrieve each spelling. The accuracy of children's performance when they were prohibited from using explicit backup strategies provided an independent estimate of the difficulty of spelling each word, which was important for assessing the adaptiveness of strategy use. When the difficulty of items is estimated from performance under free-choice conditions, the estimated difficulty of each item is confounded by two types of selection effects: selection of strategies for each word and selection of participants who use each strategy on a given word. A difficult item might appear easier than it actually is, because frequent use of effective backup strategies would improve accuracy on it, and because only the most knowledgeable children use easy-to-execute strategies such as retrieval on it. Prohibiting use of explicit backup strategies reduces this problem, because children would use similar approaches to spell the word. Percent correct in this explicit-backup-strategies-prohibited condition could then be correlated with how often children used backup strategies in the free-choice condition, thus providing an estimate of how closely strategy choices were related to word difficulty.

Differences in performance between the two conditions also provided an index of the effectiveness of children's explicit backup strategy use. In parallel experiments in algorithmic domains, being allowed to use backup strategies has resulted in more accurate performance (Siegler & Lemaire, 1997; Siegler & Shragr, 1984). Although it is not possible to suppress all use of explicit backup strategies in spelling, their use should occur much less frequently when the instructions prohibit them and when the experimenter reminds children not to use them if they ever overtly violate the prohibition. Thus, if explicit backup strategies promote accurate spelling, accuracy should be higher in the backup-strategies-allowed condition.

Finally, to examine stability of individual differences, the same children were tested in both first and second grade. Previous longitudinal studies of spelling (Beers, 1980; Gentry, 1982; Gerber, 1984; Nunes, Bryant, & Bindman, 1997) have focused solely on consistency of individual differences in error quality and whether children progressed through the predicted
stages. The present design yielded data on consistency of strategy use as well as on consistency of accuracy and specific errors over the 1 year period.

Overview. We examined whether three basic phenomena posited by the overlapping waves model—widespread variability, adaptive choice, and gradual change—are present in spelling. Because of the non-algorithmic nature of spelling strategies, their contribution to accurate spelling was also a key issue. The trial-by-trial assessment of strategies allowed us to identify which strategies the first and second graders used, how often they used them, how accurately they used them, whether they chose adaptively among them, and how strategy use changed over the one-year period. We hypothesized that even first graders would use retrieval and a variety of backup strategies, that the effectiveness of backup strategies would be considerably lower than in algorithmic domains, and that children would choose adaptively among the strategies. We also predicted that improvements in performance after a year of experience would result from a combination of addition of new strategies, greater use of more effective strategies, improved execution of each strategy, and more adaptive choices among the strategies.

METHOD

Participants

Thirty students (16 female and 14 male) participated; their mean age at the beginning of the study was 6 years, 10 months. All attended a private, urban, elementary school at which they received daily spelling lessons. When tested in November of first grade, they had received less than 3 months of formal spelling instruction.

Of the 30 children who were tested in first grade, 23 (77%) were retested in November of second grade. Among them, 12 were female, 11 male; the other 7 children either had left school or did not have parental permission to participate again. First grade spelling performance was similar for children who participated in the experiment as second graders and for those who did not. The experimenter was the first author.

Materials

Two lists were constructed for the children in each grade. The lists presented to first graders contained 15 words taken from the beginning, middle, and end of a first-grade spelling book. For each word in one list, a word with a similar spelling was included in the other. For example, one list included “hat,” “fish,” and “clown”; the other included “bat,” “dish,” and “frown.”

Each list presented to second graders contained 24 words—15 new words and 9 old ones. The new words were taken from the beginning, middle, and end of a second-grade spelling book. Again, the spelling patterns of words in the two lists were matched. The nine old words were sampled from the lists presented when the children were in first grade, so that pairs of words from the two lists were of equivalent difficulty, as assessed by percent of errors on the words in first grade. The four lists are shown in the Appendix.

Design

Each child spelled words under two conditions. In the allowed condition, explicit backup strategies were permitted. Children were told to take their time and to do anything they wanted to spell each word. In the prohibited condition, children were told not to use explicit backup strategies. That is, they were told to write the first spelling that came to mind and not to do anything special, such as sounding out the word.

Half of the children participated in the allowed condition first, the others participated in the prohibited condition first. Each list of words was presented equally often in both conditions. In second grade, children were presented the nine old words on each list in the same experimental condition as in the previous year.

Procedure

Each child was tested individually in a room in the school on 2 days, approximately 4 days apart. When children arrived at the experimental room, they were presented the instructions for their first condition and an easy practice word to spell. Then the words in the list were presented in random order. On each word, children were asked to name a picture that illustrated the word, and then to spell the word. In the allowed condition, after they spelled each word, they were asked “How did you figure out how to spell ___?” On the few occasions that a child said that she did not know, she was asked “Did you just know how to spell it? Sound it out? Use another word to help you spell it? Use a rule? Do anything else?” For words with morphological endings (-es, -ing, -ed), children were probed to see if they had used the reported strategy to spell the whole word or if they had used it to spell the root word and figured out the rest another way. Both overt and self-reported strategies were recorded. When children were in the prohibited condition, the experimenter interrupted any overt attempts to use backup strategies and prompted children to begin writing the letters if they did not do so quickly.

All sessions were videotaped. Coding of both overt
and self-reported strategy use on each trial was verified by two independent coders. When overt behavior was present, strategy classifications were based on it. When overt behavior was absent or ambiguous, the verbal report was used to classify the strategy. Solution times were recorded to the nearest .1 s. Timing began when the child identified the picture that indicated the target word and ended when the child lifted her pencil from the paper.

RESULTS

Because accurate identification of children’s strategies provided the foundation for all other findings, we first present evidence regarding the reliability and validity of the strategy assessments. Then we present evidence concerning the variability of the strategies that were used, the adaptiveness of choices among them, and the effectiveness of the strategies. Because first and second graders were asked to spell different words, we analyzed their performance separately, as well as performing some comparisons between them to examine changes in performance that occurred from first to second grade. All differences were significant at \( p < .05 \), unless otherwise noted. The order of the conditions did not influence performance; therefore, analyses including this variable are not reported.

Reliability and Validity of Strategy Classifications

The strategy assessments in the allowed condition were highly reliable. Two independent raters agreed on 93% of assessments of first graders’ overt behavior and 95% of assessments of their verbal reports. The raters agreed on 97% of assessments of both overt behavior and verbal reports of second graders. The few disagreements were resolved through discussion.

To examine the validity of the strategy classifications, we analyzed the relation between self-reported strategy use and overt behavior in the allowed condition (the condition in which self-reports of strategy use were obtained). On each trial, overt and self-reported strategy use were classified into one of four dichotomous categories (Table 1). One category consisted of trials on which children showed no overt signs of backup strategy use and reported simply knowing how to spell the word. These were classified as retrieval trials. A second category consisted of trials on which children showed overt signs of backup strategy use and reported using a backup strategy, leaving little question that they had used a backup strategy on the trial. Consistent with the classifications, solution times were much faster on trials classified as retrieval than on trials classified as overt use of backup strategies, \( t(13) = 3.29 \) for first graders, \( t(11) = 5.03 \) for second graders. (To ensure meaningful estimates of performance, only children who used the relevant strategy at least three times were included in the analysis of each strategy’s speed and accuracy.) These two categories, in which self-reports were consistent with overt behavior, comprised more than 70% of trials in both grades.

In the third category, self-reports and overt behavior were inconsistent. On these trials, children could be seen using a backup strategy, but they reported retrieving the answer. Such trials were rare: 4% of trials for first graders and less than 1% for second graders. Even here, most of the self-reports appear to have been incomplete rather than wrong. Mean solution time on these trials was within half a second of that on unambiguous retrieval trials, and it was 5 s faster than on unambiguous backup strategy trials. Accuracy on these trials also was considerably closer to that on unambiguous retrieval trials than to that on

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Strategy Classifications, in First and Second grade Divided by Observed and Reported Strategy Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed Strategy</td>
<td>Reported Strategy</td>
</tr>
<tr>
<td>---------</td>
<td>------------------</td>
</tr>
<tr>
<td>First grade</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>Retrieval</td>
</tr>
<tr>
<td>Backup</td>
<td>Backup</td>
</tr>
<tr>
<td>Backup</td>
<td>Retrieval</td>
</tr>
<tr>
<td>None</td>
<td>Backup</td>
</tr>
<tr>
<td>Second grade</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>Retrieval</td>
</tr>
<tr>
<td>Backup</td>
<td>Backup</td>
</tr>
<tr>
<td>Backup</td>
<td>Retrieval</td>
</tr>
<tr>
<td>None</td>
<td>Backup</td>
</tr>
</tbody>
</table>
unambiguous backup strategy trials (Table 1). Examination of the videotapes indicated that on these trials, children often briefly began to use an overt backup strategy, usually sounding out, but then quickly wrote the remaining letters. Thus, consistent with their verbal reports, children may have retrieved much of the spelling on these trials.

The final category consisted of the 25% of trials on which children did not overtly use a backup strategy but reported using one. This combination could arise in two ways. Children could have used a backup strategy covertly, in which case the self-report would be accurate. Alternatively, children could have retrieved the spelling, in which case the self-report would be inaccurate.

These interpretations were tested by comparing accuracy and speed on these trials to accuracy and speed generated on unambiguous backup strategy trials and on unambiguous retrieval trials. If the self-reports of backup strategy use were valid, and children used backup strategies covertly on these trials, then accuracy and solution times on them should resemble those on trials where both measures indicated use of a backup strategy. However, if the self-reports of backup strategy use were inaccurate, and children used retrieval on these trials, then accuracy and solution times on them should resemble those on trials where both verbal reports and overt behavior indicated use of retrieval.

The data supported the view that the self-reports were accurate and therefore that the strategy classifications were valid. Accuracy on trials where backup strategy use was reported but not observed was much lower than accuracy on the unambiguous retrieval trials, \( t(13) = 13.90 \) for first graders, \( t(21) = 7.42 \) for second graders. In contrast, accuracy on these trials did not differ from the accuracy on unambiguous backup strategy trials.

Similarly, mean solution times on trials where backup strategies were reported but not observed were longer than on unambiguous retrieval trials, \( t(11) = 2.07, p = .06 \) for first graders, \( t(21) = 5.83 \) for second graders. As with the accuracy measure, there was no difference between trials where backup strategies were reported but not observed and unambiguous backup strategy trials for first graders. The one exception to the otherwise clear-cut pattern was that the second graders' times on the unambiguous backup strategy trials were somewhat longer than their times on the trials where children reported using backup strategies but were not observed doing so, \( t(11) = 2.69 \). This was due in part to the slowest strategy, visual checking, being particularly likely to involve overt behavior. It was also due in part to poorer spellers (as defined by percent correct in the prohibited condition) being more likely when in the allowed condition to use backup strategies overtly rather than covertly, \( r(21) = .70 \) and to have longer mean solution times, \( r(21) = .63 \).

Overall, both accuracy and solution time comparisons supported the validity of the children's verbal reports and our distinction between retrieval and explicit backup strategies. The self-reports appeared accurate in three of the four categories, comprising 97% of trials. The high frequency of covert backup strategies made it clear that basing strategy assessments solely on overt behavior would have yielded a less accurate depiction of strategy use than the present approach of relying on overt behavior when it was present and relying on verbal report when no overt behavior was evident.

Variability of Strategy Use

Range of strategies used. Our primary focus was on which strategies children used to spell individual words in the allowed condition. Children used six strategies in this condition: retrieval, sounding out, retrieve/sounding out, drawing analogies, relying on rules, and visual checking. The names of most of the strategies indicate their content quite straightforwardly. Retrieve/sound out involved retrieving part of the word's spelling and sounding out other parts. For example, a child might quickly write the first two letters of "baby," sound out the remaining letters, and then report that she knew part of "baby" but sounded out the rest. Analogizing involved reference to knowledge of another word's spelling as the basis for spelling the presented word (e.g., citing the spelling of "fish" as an explanation of one's spelling of "dish"). Rule use involved using conventional formulas such as the silent -e rule for marking long vowels or the "ing" or "ed" rule for words having those endings.\(^1\) Visual checking involved writing an approximation of the word, changing a single letter, seeing whether it "looked right," and repeating the cycle if necessary. Rule use and visual checking were always used in combination with other strategies because neither could be used to spell an entire word. However, the specific combinations of strategies that arose were used very infrequently, so the trials were collapsed and categorized as rule use or visual checking for primary analyses.

\(^1\) Overt indicators of rule use were instances in which a child's spelling suggested use of a rule and the primary strategy used for spelling the word would not have been expected to lead to that spelling (e.g., "taked" for "talked," when sounding out was used to spell the root word). Use of analogies could not be detected through overt behavior.
Table 2  Percent of Children Who Used Each Strategy at Least Once and Percent of Trials on Which Strategy Was Used on the Complete Set of Words From Each Grade

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Percent Who Used</th>
<th>Percent of Use</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First Grade</td>
<td>Second Grade</td>
</tr>
<tr>
<td>Retrieval</td>
<td>96</td>
<td>100</td>
</tr>
<tr>
<td>Sound out</td>
<td>90</td>
<td>91</td>
</tr>
<tr>
<td>Retrieve/sound out</td>
<td>57</td>
<td>70</td>
</tr>
<tr>
<td>Analogy</td>
<td>23</td>
<td>48</td>
</tr>
<tr>
<td>Rule</td>
<td>23</td>
<td>87</td>
</tr>
<tr>
<td>Visual checking</td>
<td>10</td>
<td>65</td>
</tr>
</tbody>
</table>

Every first grader used multiple strategies to spell; individuals used 2 to 5 strategies, $M = 3.0$. As shown in Table 2, almost all of them used retrieval and sounding out, and a significant minority used analogies or rules. Sounding out and retrieval were also first graders’ most frequent strategies; they used one or the other on 88% of trials (Table 2).

Second graders used the same six strategies to spell as they had as first graders. Individuals used between 2 and 6 of them, $M = 4.5$. Almost all second graders used retrieval, sounding out, and rules. Most also used retrieve/sound out and visual checking, and almost half drew analogies. As in first grade, retrieval and sounding out were the two most common strategies; second graders used one or the other on 80% of trails.

Changes in strategy use. Both the number of children who used each strategy and the frequency of use of each strategy shifted from first to second grade. These changes were apparent in children’s performance on the complete lists of words from each grade (Table 2). Individual children used a greater number of strategies in second grade than in first grade, $M = 4.5$ and $M = 3.0$, $t(22) = 4.93$. Comparisons of number of uses of each strategy in first and second grade indicated that frequency of three strategies increased: retrieval, $t(22) = 6.00$; rule use, $t(22) = 3.96$; and visual checking, $t(22) = 2.12$. One strategy, sounding out, became less frequent, $t(22) = 7.94$.

Comparing changes in performance on the nine words that were presented to children in both grades yielded similar trends in strategy use. On this constant set of words, use of retrieval increased from first to second grade, 41% versus 78%, $t(22) = 7.78$. Corresponding decreases occurred in use of sounding out, 47% versus 12%, $t(22) = 8.33$ and retrieve/sound out, 6% versus 1%, $t(22) = 3.87$.

One difference between performance on the nine words and on the larger set was that on the nine words, children used the same number of strategies in first and second grade, 2.7 versus 2.6. This was partially because second graders could retrieve the bulk of these spellings. However, it also suggests that some of the change in number of strategies used was due to the words presented to second graders being amenable to more strategies. For example, the lists presented to second graders included more words on which standard morphological rules could be applied than did the lists presented to first graders.

We also examined changes between first and second grade in the specific strategies that individual children used. Comparisons on the complete word lists revealed that children used an average of two new strategies in second grade; 22 of the 23 children who participated in both first and second grade used at least one new strategy ($range = 0–3$). More than one-half of the children added rule use and/or visual checking, and about one-third added analogy or retrieve/sound out. Half of the children also did not use at least one strategy in second grade that they had used in first grade. Retrieve/sound out was the most commonly deleted strategy.

The changes on the nine repeated words were similar, but less dramatic. When spelling the same words a year later, half of the children used at least one new strategy, and half did not use at least one strategy they had used the year before. Analogy and visual checking were the most commonly added strategies, and retrieve/sound out was the most commonly deleted strategy.

Strategy combinations. Multiple strategy use existed within trials as well. Seventy-percent of first graders and 95% of second graders used two or more strategies to spell a single word on at least one occasion. The most common combination was retrieving part of the spelling and sounding out the rest. As shown in Table 2, most children in both grades used this combination of strategies at sometime, although it was used on only about 5% of trials. Rule use and visual checking also involved the use of multiple strategies. The different combinations of strategies

Table 3  Combinations of Strategies Used within a Single Trial

<table>
<thead>
<tr>
<th>Strategy Combination</th>
<th>Percent of Use</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First Grade</td>
</tr>
<tr>
<td>Retrieve/sound out</td>
<td>5</td>
</tr>
<tr>
<td>Rule/retrieval</td>
<td></td>
</tr>
<tr>
<td>Rule/sound out</td>
<td>2</td>
</tr>
<tr>
<td>Rule/retrieval/sound out</td>
<td></td>
</tr>
<tr>
<td>Visual check/retrieval</td>
<td></td>
</tr>
<tr>
<td>Visual check/ sound out</td>
<td>1</td>
</tr>
<tr>
<td>Visual check/retrieval/sound out</td>
<td></td>
</tr>
</tbody>
</table>
used to spell a single word are presented in Table 3. These combinations accounted for 8% of trials for first graders and 17% of trials for second graders. Children in both grades flexibly combined strategies to spell a single word.

Adaptive Choice

Percent errors for each word in the prohibited condition provided a measure of word difficulty that was independent of the varied strategies that were used on the word in the allowed condition. This independent measure of word difficulty made it possible to examine how closely frequency of use of backup strategies corresponded to problem difficulty (the first sense of adaptiveness discussed in the introduction). For first graders, this measure of word difficulty proved to be highly correlated with backup strategy use. As shown in Figure 1, the more often a word was misspelled in the prohibited condition, the more often children used backup strategies on that word when it was presented in the allowed condition, $r(28) = .85$.

Second graders' strategy choices showed comparable sensitivity to word difficulty; the corresponding correlation was $r(46) = .73$ (Figure 2). The correlations did not differ between the two grades. In addition, both relations included significant quadratic components as well as linear ones: for first graders' linear component, $F(1, 27) = 64.64$, quadratic component, $F(1, 27) = 4.67$, $R^2 = .76$; for second graders' linear component, $F(1, 45) = 60.22$, quadratic component, $F(1, 45) = 10.43$, $R^2 = .61$. Presence of the quadratic component reflected percent use of backup strategies in both grades reaching an asymptote at about 80% use for words that elicited 50% or more errors.

![Figure 1](image1.png)  
**Figure 1** Relation between percent backup strategy use and problem difficulty in first grade.

![Figure 2](image2.png)  
**Figure 2** Relation between percent backup strategy use and problem difficulty in second grade.

Effectiveness of Strategies

**Accuracy and speed of strategy use.** The six strategies used by children in the allowed condition varied considerably in speed and accuracy (Table 4). Children were correct on a much higher percentage of trials when they used retrieval than when they used backup strategies: for first graders, 86% versus 22% correct, $t(19) = 13.17$; for second graders, 96% versus 57% correct, $t(20) = 7.40$. Thus, although backup strategies were used quite often by both first and second graders (61% and 38% of words, respectively), considerably smaller percentages of words spelled correctly were generated by use of backup strategies (28% and 24% of correctly spelled words for first and second graders, respectively.) Children also were much faster when they used retrieval than when they used backup strategies: for first graders, 11 s versus 17 s, $t(19) = 3.66$; for second graders, 6 s versus 11 s, $t(20) = 9.01$.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Percent Correct</th>
<th>Mean Solution Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>First Grade</td>
<td>Second Grade</td>
</tr>
<tr>
<td>Retrieval</td>
<td>86</td>
<td>96</td>
</tr>
<tr>
<td>Sound out</td>
<td>15</td>
<td>34</td>
</tr>
<tr>
<td>Retrieve/sound out</td>
<td>57</td>
<td>61</td>
</tr>
<tr>
<td>Analogy</td>
<td>79</td>
<td>88</td>
</tr>
<tr>
<td>Rule</td>
<td>36</td>
<td>66</td>
</tr>
<tr>
<td>Visual checking</td>
<td>17</td>
<td>63</td>
</tr>
</tbody>
</table>

Table 4 Accuracy and Speed of Each Strategy on the Complete Set of Words From Each Grade
The backup strategies also varied considerably among themselves in accuracy and speed. Sounding out was the least effective backup strategy; it led to the correct spelling on only 15% of uses by first graders, 34% by second graders. The other backup strategies were more accurate than sounding out, though less accurate than retrieval, in both first grade (56% correct) and in second grade (68% correct). However, all of these other backup strategies were used rather infrequently; together, they were used on 11% of trials in first grade, 20% in second grade.

The limited effectiveness of backup strategies cannot be attributed solely to the irregularity of English orthography. Lack of skill at executing the primary backup strategy, sounding out, also contributed. The goal of sounding out is to produce a phonemic spelling, that is, a spelling in which every sound in the word is represented with a string of letters that has a reasonable probability of representing that sound in some English spelling. For example, spelling "candy" as "kandee" is phonemic, though incorrect. When the pronunciation of a word suggested a different letter than the one used in the standard spelling, the letter representing the spoken sound was also considered phonemically plausible (e.g., "t" or "d" in "letter" and "walked"). For shwa vowels that are arguably not pronounced (Treiman, 1993), no vowel was required for a spelling to be classified as phonemic (e.g., the "i" in "girl" and the second "e" in "seven"). However, even given the liberal coding system for phonemic correctness, only 50% of first graders' errors and 68% of second graders' errors on trials where they used the sounding-out strategy were phonemic. In many of the other misspellings, children did not represent all of the sounds in the word.

Changes in accuracy and speed. The speed and accuracy of strategy execution improved greatly between first and second grades. Despite being asked to spell more difficult words, second graders were faster and more accurate than first graders at both retrieval and backup strategy use: \( t(17) = 9.27 \) for speed of retrieval; \( t(17) = 3.05 \) for accuracy of retrieval; \( t(20) = 4.75 \) for speed of backup strategies; \( t(20) = 7.60 \) for accuracy of backup strategies.

Results were similar for the nine words that children were presented in both grades. Second graders were faster, 5 s versus 10 s, \( t(12) = 12.73 \), and more accurate, 99% versus 90% correct, \( t(12) = 5.48 \), on retrieval trials with these words. They were also faster, 9 s versus 16 s, and more accurate, 58% versus 27% correct, on backup strategy trials with these words. The statistical significance of changes in the speed and accuracy of backup strategies on these words could not be evaluated, because second graders rarely used backup strategies on them.

Individual differences. Each child's base spelling ability was indexed by the child's percent correct in the prohibited condition. First graders varied greatly on this measure of spelling skill; their scores ranged from 13% to 93% correct. These differences in first graders' performance were related to the strategies that children used in the allowed condition and to how effectively they executed the strategies. Better spellers used backup strategies less often; percent correct in the prohibited condition correlated \( r(28) = -.69 \) with percent use of backup strategies in the allowed condition. Better spellers also executed the backup strategies more effectively when they did use them. Percent correct in the prohibited condition correlated \( r(26) = .56 \) with percent correct on backup strategy trials in the allowed condition.

Large individual differences also were seen for the children when they were second graders. Percent correct varied from 29% to 100%. As with the first graders, second grader's percent correct in the prohibited condition was negatively related to the frequency of the child's backup strategy use in the allowed condition, \( r(21) = -.74 \), and was positively related to the accuracy of the child's backup strategy use in the allowed condition, \( r(19) = .62 \). These individual differences were not related to gender in either grade. There were no differences between boys' and girls' performance on any measure.

The individual differences in spelling skill were quite stable from first to second grade. Percent correct in the prohibited condition in first grade correlated \( r(21) = .64 \) with percent correct in that condition in second grade. Percent correct in the allowed condition correlated \( r(21) = .72 \) at the two times.

Strategy use was also quite stable over the one year period. Children who used retrieval more often in first grade also used it more often in second grade, \( r(21) = .70 \). Percent use of sounding out, the only other strategy that was used sufficiently often in both grades to yield meaningful correlations, also was quite stable from first to second grade, \( r(21) = .66 \). This stability of strategy use helps to explain the stability of accuracy described in the last paragraph. Frequent use of retrieval yielded accurate performance in both grades, and frequent use of sounding out yielded inaccurate performance in both.

Between condition comparisons. To further explore the effects of allowing children to use explicit backup strategies, we compared children's performance in the allowed and prohibited conditions. Based on the assumption that explicit use of backup strategies, even nonalgorithmic ones, would increase the accu-
racy of children’s spelling, we expected percent correct in the allowed condition to be higher than that in the prohibited condition. The data did not support this expectation. First graders answered correctly on 47% of trials in the allowed condition, 46% in the prohibited condition. Second graders answered correctly on 79% of trials in the allowed condition, 78% in the prohibited condition.

Although children did not spell more words correctly in the allowed condition, first graders’ errors were of better quality in this condition. Misspellings in the allowed condition more often had the correct number of letters than did misspellings in the prohibited condition, 30% versus 23%, t(27) = 2.28. A more comprehensive error quality score was obtained for each misspelled word by dividing the number of correct letters in the correct sequence by the number of letters in the correct spelling of the word. This error quality measure was similar to that used by Turner and Quinn (1986). For example, “bed” spelled “bad” was scored as 2/3, or .67. (See Table 5 for illustrations of error quality scores within this system.) This measure again indicated that first graders’ error quality in the allowed condition was higher than in the prohibited condition (.67 versus .63), t(25) = 2.69. Thus, although being permitted to use explicit backup strategies did not lead to first graders generating a greater number of correct spellings, it did lead to their making higher quality errors.

Unlike in first grade, second graders’ spelling errors in the allowed condition were not more likely to be of the correct length. In each condition, 25% of errors had the correct number of letters. The error quality scores for each word also did not differ between conditions (.76 versus .77). Thus, being allowed to freely use explicit backup strategies did not improve the quality of second graders’ errors.

Might the lack of difference in accuracy between the prohibited and allowed conditions, and the lack of difference in error quality among second graders, have been due to children not following the experimenter’s instructions and using the same strategies in the prohibited condition as in the allowed condition? Consistent with this view, on 10% of first graders’ and 8% of second graders’ trials, the experimenter had to interrupt overt attempts to use backup strategies in the prohibited condition.

Other data, however, argued against the view that the lack of difference in accuracy between the two conditions stemmed largely from children not following the instructions. Performance was quicker in the prohibited condition than in the allowed condition: for first graders, 14 s versus 11 s, t(29) = 3.23; for second graders, 7.6 s versus 6.0 s, t(22) = 6.24. In contrast, in both grades, mean solution times in the prohibited condition were virtually identical to mean solution times for retrieval trials in the allowed condition (11 s for both in first grade; 6 s for both in second grade). The fact that times were longer in the allowed condition than in the prohibited condition was not attributable to children in the allowed condition including more letters in their spellings. The mean solution time per letter written was also higher in the allowed condition than in the prohibited condition: for first graders, 3.9 s versus 3.4 s, t(29) = 2.21; for second graders, 1.7 s versus 1.3 s, t(22) = 6.31. Further, removing the trials on which children had to be stopped from using overt strategies did not alter the overall pattern of results. Thus, the lack of difference in percent correct between the prohibited and allowed conditions was not attributable to the distribution of strategies being the same in the two conditions.

DISCUSSION

In this study, we examined the applicability of the overlapping waves model to the development of spelling. Spelling was of particular interest because it was a nonalgorithmic domain, one in which even perfect execution of the available backup strategies did not guarantee correct answers. The study yielded more precise and accurate descriptions of the devel-

---

Table 5   Examples of Error Quality Scoring System from Children’s Misspellings of “Dish”

<table>
<thead>
<tr>
<th>Allowed Condition</th>
<th>Prohibited Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spelling</td>
<td>Score</td>
</tr>
<tr>
<td>dis</td>
<td>3/4</td>
</tr>
<tr>
<td>dash</td>
<td>3/4</td>
</tr>
<tr>
<td>doh</td>
<td>3/4</td>
</tr>
<tr>
<td>dh</td>
<td>2/4</td>
</tr>
<tr>
<td>det</td>
<td>1/4</td>
</tr>
</tbody>
</table>

Note: Score = number of correct letters in correct sequence/total number of letters in the word.

3 Of the 26 words that were misspelled by at least three first graders in each condition, error quality scores in the allowed condition were higher than those in the prohibited condition on 16 words, equal on 3, and lower on 7. Error quality scores on individual words ranged from .44-.78.

4 Of the 13 words that were misspelled by at least three second graders in each condition, error quality scores in the allowed condition were higher than those in the prohibited condition on 7 words, the same on 2, and lower on 4. Error quality scores for individual words ranged from .70-.87, with the exception of one word that had a mean error quality of .39.
opment of spelling than were previously available. It also yielded a new method for analyzing the validity of trial-by-trial strategy assessments and several theoretical implications regarding strategy choice in nonalgorithmic domains. In this concluding section, we consider each of these contributions in turn.

The Development of Spelling Skills

In the present study, we examined first and second graders while they were in the process of spelling words. This on-line examination of spelling yielded a very different depiction than that yielded by previous post hoc attempts to infer knowledge of spelling from corpuses of written material. The differences can be summarized within the three main themes of the overlapping waves model: (1) abundant variability, (2) adaptive choice, and (3) gradual change.

First consider findings regarding variability. Previous stage-oriented depictions have posited that beginning spellers (the first graders in the current study) rely on sounding out, whereas those at later stages (the second graders in the current study) also use a variety of auxiliary strategies, such as relying on rules and drawing analogies (Ehri, 1992; Gentry 1982; Henderson, 1980). The present findings, in contrast, suggested that spelling strategies were variable from early in learning. Rather than first graders’ spelling fitting into one stage and second graders’ into another, the same six strategies were used at both ages and the same two strategies accounted for more than 80% of children’s spellings at each age. Substantial variability was present within individual children’s spelling at each time of measurement; in first grade, children used an average of three strategies, and in second grade, they used an average of more than four strategies. This variability also was present at times within a trial. On individual trials, both first and second graders flexibly combined retrieval, sounding out, rule use and visual checking.

Many of these observations would have been impossible given the usual method of examining corpuses of previously written material. The present online observations allowed identification of a greater variety of strategies, assessment of how often each was used, and specification of combinations of strategies that were used on individual trials. Variability in strategy use was present at all levels.

Considerable variability was also present between children. Individual first and second graders differed widely in spelling skill and strategy use. Better spellers used backup strategies less often and executed the backup strategies more effectively when they did use them. Better spellers in first grade remained better spellers in second grade. Stability in the frequency of use of retrieval and sounding out partially accounted for the stable individual differences in speed and accuracy.

Because the present method allowed assessment of strategy use on each trial, it also allowed us to examine the adaptiveness of choices among the strategies. This issue has not been systematically studied in past studies of spelling. Both first and second graders’ strategy choices were highly adaptive, in the sense that the more difficult the word, the more often that children used a backup strategy. On the other hand, they were not very adaptive in the sense that the choices did not lead to greater accuracy than was possible without use of explicit backup strategies. Being able to use explicit backup strategies led to some improvement in the quality of first graders’ errors, but this was not observed with second graders. Nonetheless, the correlations between problem difficulty and use of backup strategies—r = .85 and r = .73 for first and second graders, respectively—were comparable to those that have been observed in domains where backup strategies are much more effective, such as addition and multiplication (Lemaire & Siegler, 1995; Siegler & Robinson, 1982). The results suggest that children may have based decisions to use backup strategies largely on the likelihood of retrieval not yielding a correct answer, rather than on the likelihood that backup strategies would be more effective than retrieval.

Finally, the results identified several contributors to improvements in speed and accuracy from first to second grade, as well as indicating that several plausible sources of improvement did not contribute. The improvements in speed and accuracy appeared due to three more specific changes. One was increasing use of the fastest and most accurate strategy, retrieval. Another was movement away from the slowest and least accurate backup strategy, sounding out, and toward faster and more accurate backup strategies, such as drawing analogies and relying on rules. The third source of improvement was increasingly quick and accurate execution of all strategies.

Two other factors that might have been expected to contribute to faster and more accurate performance did not. One was introduction of new strategies, the factor emphasized in stage approaches. The same six strategies were used in first and second grade. Although most children added some strategies between first and second grade, these new strategies were used rather infrequently in both grades. It is possible that on a broader set of words, the new strategies would have been used more often. However, the ap-
plicability of the new strategies that were added, in particular drawing analogies and applying rules, will always be limited, and strategies such as rule use are only helpful for spelling part of each word to which they apply. Addition of new strategies seems to reflect a more sophisticated understanding of English orthography, but their use appears to have little impact on overall speed and accuracy, at least initially.

Another potential source of improvement was increasingly adaptive choices among strategies. However, choices in the two grades proved equally sensitive to problem difficulty. Thus, trial-by-trial examination of strategy use allowed us to identify factors that contributed to the global improvements in spelling between first and second grade, and also factors that did not contribute to them.

Methods for Validating Strategy Assessments

The current approach to studying development rests on a foundation of reliable and valid assessments of strategies on each trial. If strategy assessments lack these qualities, the research will not be informative. Fortunately, in the present study (as in previous ones that used the current strategy assessment method), assessments proved to be reliable and valid. The method is to base strategy classifications on overt behavior when that indicates which strategy was used and to base classifications on immediately retrospective self reports when overt behavior is absent or ambiguous. As in previous studies, the strategies identified by this method differed in solution time and accuracy characteristics in ways that were consistent with the inherent nature of the strategies.

The present study added a new kind of validation, one that focused on the verbal reports. If verbal reports of strategy use are accurate, speed and accuracy should be similar on trials where the verbal report provides the only evidence for use of a strategy and on trials where both verbal reports and overt behavior attest to its use.

The verbal reports of spelling strategies passed this test. Speed and accuracy were similar on trials where only the verbal report indicated use of a backup strategy to that on trials where both overt behavior and verbal reports indicated its use. Even on the very small percentage of trials on which the verbal report was inconsistent with the overt behavior, solution time and accuracy data suggested that the verbal report was incomplete rather than wrong. Thus, children's immediately retrospective verbal reports appeared to accurately reflect their spelling strategies.

Strategy Choice in Nonalgorithmic Domains

The present study was the first detailed comparison of strategy choices in a nonalgorithmic domain to choices made in algorithmic domains. The algorithmic/nonalgorithmic distinction was of theoretical interest because the advantage of using backup strategies was expected to be lower in nonalgorithmic domains. We thought that this reduced advantage might reduce the usual close correspondence between problem difficulty and frequency of use of backup strategies.

The advantage of being allowed to use backup strategies was indeed lower than in previously studied domains. Even 4- and 5-year-olds, just beginning to add and subtract, use backup strategies quite accurately. They generate about 75% correct answers on problems with sums of 10 or less when they use backup strategies (Siegler & Robinson, 1982; Siegler & Shrager, 1984). In contrast, spelling backup strategies yielded much lower accuracy. Backup strategies only led to a correct spelling on 22% of uses in first grade and 57% of uses in second grade, compared to accuracy levels of 86% and 96% for retrieval in first and second grade, respectively. Further, the contribution of backup strategies to overall spelling accuracy was limited; only one quarter of the correct spellings in the allowed condition were produced by use of an explicit backup strategy in either grade. Finally and most compelling, in contrast to previous research in algorithmic domains, spelling accuracy of both first and second graders was unaffected by whether use of explicit backup strategies was allowed or prohibited. Although it may not be possible to suppress all use of explicit backup strategies, the large differences between times to spell words in the two conditions, and in the amounts of overt behavior that were observed, suggests that the prohibition of backup strategies was at least somewhat effective. The almost identical percentages correct in the allowed and prohibited conditions at both ages suggests that explicit backup strategies have minimal impact on spelling accuracy. For beginning spellers, use of rapid, automatic processing seems to produce as many correct spellings as use of explicit, time-consuming, backup strategies.

Under these conditions, children might have been expected to use explicit backup strategies rarely if at all. If backup strategies take longer than retrieval and have little impact on accuracy, why use them? Yet children not only used them on a substantial percentage of trials, they also chose among them very systematically. The pattern of strategy choice closely resembled that observed previously in algorithmic domains, where being allowed to use backup strate-
gies substantially improves performance. How can this be explained?

In algorithmic domains, children's strategy choices balance considerations of speed and accuracy. They tend to use the fastest strategies that are likely to yield the correct solution. The more favorable the speed and accuracy yielded by backup strategies relative to those yielded by retrieval, the more likely that backup strategies will be chosen (Siegler & Lemaire, 1997). This suggests that strategy choices are based on the degree of difference between the outcomes of alternative strategies. The choice mechanism postulated by recent models of strategy choice works in this way (Shrager & Siegler, 1998; Siegler & Shipley, 1995).

However, results from the present study suggest a simpler interpretation of adaptive strategy choice, at least in nonalgorithmic domains such as spelling. Increased use of backup strategies on difficult problems may reflect a tendency to do something plausible when retrieval seems unlikely to yield the correct solution. Within this interpretation, children use retrieval whenever they are sufficiently confident in the correctness of the answer. If the associated strength of the answer falls below the confidence criterion, they use a backup strategy, regardless of its likelihood of success. This interpretation is consistent with the strategy choice mechanism postulated in Siegler and Shrager's (1984) computer simulation of strategy choices in addition. Children's strategy use in spelling indicates that the relative advantage of available strategies for speed and accuracy is not the only mechanism driving strategy choice, and that determinants of choice may vary from domain to domain.4

Regardless of the mechanisms that generate strategy choices in different domains, it is still surprising that beginning spellers use explicit backup strategies so often when they yield minimal benefits for accuracy. The phenomenon of children using strategies that do not initially increase their percentage of correct answers is not limited to spelling. Labeled a "utilization deficiency" by Miller (1990), the pattern has been observed in many studies of the development of memory and selective attention (for reviews of this literature, see Bjorklund, Miller, Coyle, & Slawinski, 1997; Miller & Seier, 1994). The present study demonstrates the pattern in a new domain—spelling—and documents an intriguing new correlate—systematic choices of when to use the strategies, despite their poor effectiveness. Results from this study also suggest that closer approximation to the correct solution is a potential motivation for using strategies that initially do not improve accuracy. Continuous measures of performance, such as answer quality, offer a more sensitive means for detecting improvements than do categorical classifications of answers as correct or incorrect. Elementary school children can distinguish between high and low quality spelling errors (Gerber, 1984), so improvements in error quality are a possible source of first graders' use of spelling backup strategies, despite the strategies not improving accuracy. The lack of a comparable effect among second graders make this interpretation uncertain, but improvements in error quality represent a plausible explanation as to why children persist in using strategies that do not lead to improved accuracy.

Previous research on utilization deficiencies also has implications for spelling. On memory and attentional tasks, initially ineffective strategies eventually improve performance. The same may be true for spelling. In the present sample, the accuracy of backup strategies in spelling increased from first to second grade, but it remained well below the limits imposed by English orthography. With further improvement in the effectiveness of backup strategies, they may come to yield more accurate performance than would otherwise be possible. Indeed, recent research indicates that backup strategies become substantially more effective from second to fifth grade (Steffler et al., 1998).

The present research indicates that the overlapping waves model can be applied to nonalgorithmic domains, as well as to algorithmic ones. The abundant variability, adaptive choice, and gradual change previously observed in arithmetic, time telling, and other algorithmic domains are all present in spelling as well. On the other hand, use of explicit backup strategies seemed to have far less of a positive effect on accuracy in spelling than in previously studied algorithmic domains. This finding raises the question of why children persist in using time-consuming backup strategies that initially do little to improve performance. Answering this question may help us better understand how children arrive at their strategy choices, as well as differences in learning processes in algorithmic and nonalgorithmic domains.

4Strategy choices in arithmetic may initially be based on a similar mechanism. It may be that some backup strategies in arithmetic do not improve accuracy at the very outset of their use, and that children's strategy choice is initially based on the speed and accuracy of retrieval alone. In opposition to this view, there is some evidence that children's arithmetic backup strategies are fairly effective from the very beginning; Saxe, Guberman, and Gearhart (1987) found that virtually as soon as children begin to use addition strategies (other than guessing), the strategies are fairly accurate. Lemaire and Siegler (1995) obtained similar results with multiplication. Being able to use backup strategies helped children who were in the first week of learning multiplication.
REFERENCES


ACKNOWLEDGMENTS

This research was supported by grants from the National Institutes of Health (HD-19011), National Institutes of Mental Health Training Grant (MH 19102), and a National Science Foundation graduate fellowship. We would like to thank the children, teachers and administrators of Winchester Thurston School, who made the project possible, as well as Alice Siegler and Melissa Ippolito for their help coding data. We would also like to thank Martha Alibali, David Klahr, Sharon Carver, and three anonymous reviewers for providing particularly helpful comments on previous versions of the manuscript. Portions of this research were presented at the American Educational Research Association annual meeting, April, 1996, New York, NY.

ADDRESSES AND AFFILIATIONS

Corresponding author: Bethany Rittle-Johnson, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213; e-mail: b2e@andrew.cmu.edu. Robert S. Siegler is also at Carnegie Mellon University.

APPENDIX

<table>
<thead>
<tr>
<th>First grade word lists</th>
<th>Second grade word lists</th>
</tr>
</thead>
<tbody>
<tr>
<td>List 1</td>
<td>List 2</td>
</tr>
<tr>
<td>hat</td>
<td>bat</td>
</tr>
<tr>
<td>bug</td>
<td>rug</td>
</tr>
<tr>
<td>dog</td>
<td>pop</td>
</tr>
<tr>
<td>red</td>
<td>bed</td>
</tr>
<tr>
<td>kite</td>
<td>dime</td>
</tr>
<tr>
<td>stop</td>
<td>hand</td>
</tr>
<tr>
<td>fish</td>
<td>dish</td>
</tr>
<tr>
<td>feet</td>
<td>leaf</td>
</tr>
<tr>
<td>candy</td>
<td>baby</td>
</tr>
<tr>
<td>boat</td>
<td>goat</td>
</tr>
<tr>
<td>moon</td>
<td>spoon</td>
</tr>
<tr>
<td>chair</td>
<td>stairs</td>
</tr>
<tr>
<td>girl</td>
<td>bird</td>
</tr>
<tr>
<td>letter</td>
<td>water</td>
</tr>
<tr>
<td>brown</td>
<td>clown</td>
</tr>
<tr>
<td>bug</td>
<td>red</td>
</tr>
<tr>
<td>moon</td>
<td>dish</td>
</tr>
<tr>
<td>feet</td>
<td>dime</td>
</tr>
<tr>
<td>brown</td>
<td>spoon</td>
</tr>
</tbody>
</table>


