Journal of Experimental Psychology:
Human Perception and Performance
2010, Vol. ●●, No. ●, 000–000

0096-1523/10/$12.00 DOI: 10.1037/a0019258

Interpreting Chicken-Scratch: Lexical Access for Handwritten Words

Anthony S. Barnhart and Stephen D. Goldinger
Arizona State University

Handwritten word recognition is a field of study that has largely been neglected in the psychological literature, despite its prevalence in society. Whereas studies of spoken word recognition almost exclusively employ natural, human voices as stimuli, studies of visual word recognition use synthetic typefaces, thus simplifying the process of word recognition. The current study examined the effects of handwriting on a series of lexical variables thought to influence bottom-up and top-down processing, including word frequency, regularity, bidirectional consistency, and imageability. The results suggest that the natural physical ambiguity of handwritten stimuli forces a greater reliance on top-down processes, because almost all effects were magnified, relative to conditions with computer print. These findings suggest that processes of word perception naturally adapt to handwriting, compensating for physical ambiguity by increasing top-down feedback.

AQ: 1 *Keywords:*

According to an article in the *Washington Post* (Pressler, 2006), elementary education in penmanship is disappearing in the United States, because computers are increasingly prominent in the classroom. Students are declining in the ability to read and write handwritten words, which may have unforeseen consequences, because handwriting acquisition and fluency have been shown to improve writing and composition skills (Graham, Harris, & Fink, 2000). Cursive handwriting originated in the Middle Ages, when it was used by monks to speed the process of copying religious texts; its form reflects the limitations imposed by archaic writing instruments (Lorette, 1999). Technological advances of the 20th century have dramatically changed communication: word processing, e-mail, and text messaging have arguably made synthetic typescript the predominant form of written communication. Although these changes have doubtless increased the efficiency of communication, some evidence suggests that natural handwriting conveys more information than typescript. While no empirical evidence supports the grandiose claims of graphology (i.e., that personality traits are detectable from handwriting styles; Furnham & Gunter, 1987; Neter & Ben-Shakhar, 1989), King and Koehler (2000) reported that "gender, socioeconomic status, and degree of literacy" can be inferred reliably from handwriting (p. 336). Loewenthal (1975) found that people could manipulate their handwriting to convey specific personality traits to readers. Thus, handwriting potentially expresses additional information to readers, relative to automated text.

Beyond potentially expressing indexical information, such as a writer's gender or mood, handwriting presents interesting perceptual challenges. In a manner similar to phonemes in speech (Pisoni & Luce, 1987), the individual segments of handwriting are "noisy" and nonuniform, displaying context-conditioned variation and, most likely, some degree of random variation. Computer-generated letters are composed of clear featural information; it is theoretically conceivable that computer-generated words are recognized based entirely on "bottom-up" assembly processes. We of course know that word perception is not so simple: Given pristine characters, peoples' performance is still affected by many lexical variables, such as word frequency, semantic predictability, spelling-sound consistency, neighborhood size, and others (e.g., Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Perry, Ziegler, & Zorzi, 2007). Although these effects have different empirical profiles and potential theoretical origins, they may all be classified as "top-down" effects, because word-level knowledge modulates the process of converting letter strings into perceptual objects. A natural question is whether such lexical effects might be enhanced if bottom-up perceptual processes were less robust. Because handwritten letters have greater potential for featural ambiguity, they should more often require disambiguation from word-level context. In this article, we report four experiments comparing the effects of well-known lexical variables between two different forms of input, computer print and human cursive.

In conceiving the present study, we found that the word perception literature shows a striking asymmetry across modalities. Specifically, in research on spoken word perception, the vast majority of published experiments use stimulus items that are recorded by human volunteers. Although research on "low-level" speech perception (i.e., phoneme perception) often uses synthetic materials, studies focused on lexical variables almost exclusively use naturally produced items (Duffy & Pisoni, 1992). As a result, segmental information across words is highly variable, even within experiments, and the same nominal words are potentially quite

Anthony S. Barnhart and Stephen D. Goldinger, Department of Psychology, Arizona State University, Tempe, Arizona.

This work was supported by National Institutes of Health (NIH)/National Institute of Child Health and Development (NIDCD) Grant R01-DC04535-10 to Stephen D. Goldinger. We thank Megan Papesh for generating stimulus materials. We also thank Kathleen Powers and Hayley Simkins for help in data collection, and Drs. Guy Van Orden and Michael Cortese for their comments on a previous version.

Correspondence concerning this article should be addressed to Stephen D. Goldinger, Department of Psychology, Box 871104, Arizona State University, Tempe, AZ 85287-1104. E-mail: goldinger@asu.edu
AQ: 4

tapraid5/zfn-xhp/zfn-xhp/zfn00310/zfn2492d10z | xppws | S=1 | 3/20/10 | 11:05 | Art: 2008-1190

different from one study to the next. By contrast, the literature on printed word perception reveals precisely the opposite pattern: Although the literature is vast, there is a near-total lack of research with naturally-produced materials. In fact, to our knowledge, only one prior study (by Manso De Zuniga, Humphreys, & Evett, 1991) has examined the effects of lexical variables in handwritten word-reading. To the degree that prior research has examined handwriting, it has focused on means of normalizing ambiguous physical attributes of the stimuli. Such research is not typically geared toward understanding human word perception, but to help develop handwriting recognition software.

As noted by Manso De Zuniga et al. (1991), the use of pristine stimuli may provide an incomplete understanding of the processes involved in human word recognition. Computers can easily interpret typewritten text, but programming is far more complex for recognizing handwriting. Lorette (1999) noted two major difficulties posed by handwriting: *polysemy* and *segmentation*. Polysemy is essentially ambiguity: The same nominal handwritten characters change their physical forms across contexts, and very similar forms may signal different intended characters in different contexts. Segmentation denotes the problem of determining where one character ends and another begins. These problems are central to speech perception, wherein a phoneme's context is often just as important as the phoneme itself for perception (Fowler, 2005; Goldinger, Pisoni, & Luce, 1996; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). In similar fashion to speech perception (Pisoni & Luce, 1987; Samuel, 1981), ambiguity in handwriting should naturally force readers to rely more heavily on top-down processing. If so, lexical variables might be expected to have larger effects when people process handwritten words, relative to typewritten words. As a concrete example, consider the implications of handwriting in the classic interactive-activation (IA) model (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). In the IA model (and its descendants, e.g., Grainger & Jacobs, 1996), the connections between visual features and letter nodes were set to very high weights, such that feature combinations activated correct letters quickly and perfectly. In addition, inhibition strengths between feature and letter nodes far outweighed excitation strengths. Thus, with noisy input such as handwriting, the system would likely be unable to settle at the letter level because net inhibition from inconsistent features would prevent adequate excitation. In order to overcome this, activation at the word level would have to feed back to the letter level, disambiguating the input based on lexical knowledge. Thus, processing should be heavily reliant on top-down influences, as can be measured through the manipulation of lexical variables such as word frequency.

In an early study of handwritten word perception, Corcoran and Rouse (1970) presented participants with handwritten and typed words in either blocked or mixed lists. By comparing these conditions with blocked and mixed recognition trials with uppercase and lowercase typed words, they concluded that two different "recognition programs" were involved in the processing of human- and computer-generated word forms: Recognition of handwritten words only showed a decrement when they were mixed with typed words. This finding was later disputed by Manso De Zuniga et al. (1991), who described four experiments on handwritten word perception. Their first experiment assessed repetition effects to handwritten and typewritten words, using a procedure with alternating naming and lexical decision trials. In some instances, the word in a naming trial was identical to the word in the subsequent lexical decision trial (the *immediate repeat* condition). In others, the naming word appeared as a lexical decision stimulus 15 trials later (the *lagged repeat* condition). Handwritten and typewritten stimuli either acted as primes for themselves or as primes for their counterparts. The authors reasoned that, if lexical activation occurs irrespective of script, then priming should occur equally, regardless of changes in form. Conversely, if lexical access is affected by surface form, then repetition effects should differ based on forms and lag. In the immediate repeat condition, stronger priming was observed for same-type stimuli. In the lagged repeat condition, there was no script correspondence effect, but repetition effects were stronger for all handwritten targets. These results suggested that surface information about the primes faded quickly from memory, but that priming benefits are generally stronger for handwritten words, presumably because their perception requires greater perceptual processing.

In their second experiment, Manso De Zuniga et al. (1991) had participants make lexical decisions to handwritten and typewritten stimuli, presented either clearly or with reduced contrast. Although they expected the contrast reduction to selectively impair perception of handwritten stimuli, no such interaction appeared. The authors suggested that handwriting and contrast reduction influence different stages of word processing. This hypothesis was supported in their third and fourth experiments. They again used the alternating naming / lexical decision paradigm, this time with low- and high-frequency words. In Experiment 3, script forms were manipulated between subjects; in Experiment 4, participants viewed both script forms. Both experiments produced the same results, with reliable effects of frequency and script and marginal interactions. Although statistically weak, the interactions suggested that a general "input clean-up" occurs equally in an early processing stage for all word forms, regardless of physical ambiguity, and that, counter to Corcoran and Rouse's (1970) assertion, handwriting affects the subsequent rate of lexical access.

In an experiment similar to the lagged-repeat condition in Experiment 1 of Manso De Zuniga et al. (1991); Brown and Carr (1993) examined repetition effects in lexical decision and naming for items varying in surface form (from typeface to cursive). For printed items, they observed significant repetition priming, regardless of the script of the initial exposure stimulus. For the handwritten words, the size of the priming effect was dependent on the script of the first exposure. There was a large benefit for cursive-cursive repetitions, relative to printed-cursive repetitions, suggesting that people learn to recognize the consistencies in a script over time in a fashion similar to the *font tuning* effect reported by Walker (2008).

In the present study, we extended the research by Manso De Zuniga et al. (1991), exploring various lexical effects across synthetic and natural word tokens. Our general expectations were simple: Given less familiar and predictable stimuli, we expected slower overall reading, which would increase the effects of most lexical variables. And, given the need for greater word-level interpretation of ambiguous input, we expected especially profound increases for effects that reflect top-down processing, such as word frequency effects. Across four experiments, we examined variables that seem naturally divided into "feed-forward" and "feedback" categories. We tested effects of spelling-sound regularity

tapraid5/zfn-xhp/zfn-xhp/zfn00310/zfn2492d10z | xppws | S=1 | 3/20/10 | 11:05 | Art: 2008-1190

and consistency, both relating to the feed-forward process of converting print to sound. Although these effects might be increased in handwriting (because of general slowing), we did not anticipate profound increases. By contrast, we also tested effects of word frequency, feedback consistency, and semantic imageability. These variables are all theoretically related to top-down processing (e.g., Harm & Seidenberg, 2004; McClelland & Rumelhart, 1981; Perry et al., 2007; Van Orden & Goldinger, 1994), as word-level knowledge is hypothesized to accelerate or delay the perceptual resolution of letter strings.

Experiment 1A: Frequency and Regularity Effects in Naming

One of the complications to learning English is its widespread irregularity in spelling-sound correspondences. Certain writing systems, such as Japanese kana, have a perfect, one-to-one mapping of spelling patterns to phonology (Seidenberg, Waters, Barnes, & Tanenhaus, 1984). By contrast, many English words have spelling-sound correspondences that cannot be predicted by rules of orthographic to phonologic (O-P) translation, and many lexical neighborhoods contain irregularities. Glushko (1979) argued that irregularity does not only reflect a word's adherence to O-P rules, but also the consistency of its pronunciation, relative to similar words. By this view, regularity has two levels. A "regular" word has a pronunciation that is easily determined through rules of O-P translation *and* has a consistent neighborhood (e.g., NAME follows the O-P rule that "a" is long if the word body ends in "e" and it has no conflicting neighbors). A word can also be "regular-inconsistent," meaning that its pronunciation is regular based on its spelling, but that at least one neighbor has a different pronunciation (e.g., GAVE is regular-inconsistent because its neighbor, HAVE, is an exception word). Indeed, Glushko (1979) found that both regular-inconsistent and exception words produced significantly longer naming latencies than regular-consistent words, indicating that neighborhood consistency affects word processing.

Subsequent research by Seidenberg et al. (1984) suggested that Glushko (1979) may have overestimated the effects of regular-inconsistent words on naming, given an artifact of his design. In Glushko's study, many word bodies were repeated across regular-inconsistent and exception words. Thus, longer response times (RTs) for regular-inconsistent words may have emerged because participants encountered pronunciation bodies that they had already spoken differently in earlier trials. Seidenberg et al. tested this by manipulating the order in which Glushko's words were presented. When regular-inconsistent words were presented after exception words, with comparable word bodies, they were named relatively slowly. This effect disappeared when the order was reversed, although RTs to exception words were always slow.

Word frequency influences naming time across nearly all experiments, and Manso De Zuniga et al. (1991) found increased frequency effects with handwritten items. In Experiment 1, we focused on the interaction of frequency with regularity: Seidenberg et al. (1984) proposed that high-frequency words should not have to pass through the same O-P translation as less common words because their familiar forms can directly activate pronunciations stored in memory (see Coltheart, Curtis, Atkins, & Haller, 1993). They collected naming RTs to high- and low-frequency regular and exception words. Consistent with the dual-route hypothesis, they found a regularity effect (slower naming for exception words) only for low-frequency words. Jared (2002) later assessed the same prediction using a better-controlled stimulus set. While still orthogonally manipulating word frequency and regularity, Jared controlled many other variables, including neighborhood consistency, length, initial letters and phonemes. With these controls in place, the Frequency $\times$ Regularity interaction was reduced (as were regularity effects, themselves), suggesting that regularity has little effect beyond neighborhood consistency. In Experiment 1, we sought to replicate and extend these findings, using the stimuli from Jared (2002; Experiment 2) in different formats for different groups of participants. The words were presented either in computer print (a direct replication with uniform, familiar forms) or human cursive (highly nonuniform and unfamiliar).[1] As noted earlier, our first prediction was that word frequency effects would increase, given less reliable visual features in the human cursive condition. Extending the dual-route logic from Seidenberg et al. (1984), a second prediction was that, when words are presented in cursive, regularity effects should increase among low-frequency words, and may even become reliable among high-frequency words. Deprived of familiar visual patterns, people may be unable to access the lexicon by direct access, even for high-frequency words. Fn1

Based on recent work by Balota, Yap, Cortese, and Watson (2008), all experiments in this article included standard analyses of mean RTs, and also distributional analyses. Because RT distributions tend to be positively skewed, changes in mean RTs could reflect either distributional shifting (as is often assumed) or distributional skewing, which raises the mean without changing the modal RT. Balota et al. (2008; Balota & Spieler, 1999) demonstrated that common lexical effects are often because of combinations of distributional shifting and skewing. To determine the nature of lexical effects, they proposed fitting observed RT distributions to an ex-Gaussian function, which combines the normal Gaussian distribution and an exponential distribution. This generates three parameters reflecting different aspects of the distribution. The μ and σ parameters represent the mean and *SD*, respectively, of the Gaussian portion of the distribution. Thus, these parameters (especially μ) reveal effects of distributional shifting, with effects because of outliers in the upper tail partialled out. The τ parameter represents the mean and *SD* of the exponential segment of the distribution, providing a numerical value representing distributional skewing. Using distributional analyses, Balota and Spieler (1999) found that frequency effects were the result of both shifting and skewing in the distribution. In the interest of brevity, we report only the outcomes of the ex-Gaussian analyses, which nearly always agreed with the traditional analyses, making special note of any exceptions. Standard item (F_2) analyses are presented in tables.

[1] Our original design contained four script forms, because we intended to test effects across a continuum: computer print, a computer-generated cursive, human print, and human cursive. The computer-generated cursive and human print conditions produced results that were equivalent to the computer print condition and are thus excluded for brevity.

Method

Subjects. Thirty-nine Arizona State University students participated for partial course credit (19 in the print condition and 20 in the cursive condition).

Stimuli. One hundred sixty monosyllabic words used by Jared (2002; Experiment 2) were employed in Experiment 1. These words varied on two dimensions: Frequency (high, low) and regularity (regular, exception), creating four categories, with 40 words apiece. Items were matched for neighborhood consistency, length, and initial letters and phonemes. Stimuli were duplicated in two between-groups conditions: print and cursive. Computer print words appeared in 45-point Courier New font. The volunteer writer for the human cursive condition was chosen based on the legibility and consistency of her handwriting.[2] It is difficult to find people who naturally write in a purely cursive format. The volunteer had elements of her handwriting that would fall into the "print" category, but her handwriting was predominantly cursive. Stimuli were written using a Logitech io2 digital pen: This instrument appears like a standard ball-point pen, and allows a person to write on paper, as usual. It is equipped, however, with a small camera which "reads" a fine dot pattern printed on each sheet of paper, converting this information into a digital copy of the pen strokes. The stimuli were thus generated in a digital format that could be edited and stored for later presentation as picture files. The volunteer was requested to write the stimuli in her normal hand and was given multiple opportunities to rewrite all items to ensure that they appeared legible and natural. The digitized images were sharpened and enlarged using Adobe Photoshop. All stimuli across conditions were approximately the same size (comparable to 45-point Courier New) when presented on the screen (examples of the stimuli are shown in Figure 1).

Apparatus. Participants were tested individually in a sound-attenuated booth. Stimuli were presented via the E-Prime 2.0 program (Schneider, Eschman, & Zuccolotto, 2002) on a Dell computer with a flat-panel CRT monitor, and vocal response times were collected using a standard voice key connected to an E-Prime SR Response Box.

Procedure. After informed consent was obtained, participants were acclimated to the sensitivity of the voice key through a simple naming task using written numbers from one to ten. Once ready, they proceeded immediately into the experiment. Word form was a between-subjects manipulation; each participant received words in only one script format. The experiment began with nine practice trials, intended to familiarize participants with the trial structure and presentation script. Each trial began with a fixation point appearing for 750 ms, followed by the target word. Once the participant initiated a vocal response, the word disappeared; after a 1,000-ms intertrial interval, the next trial began. If no response was detected within three seconds, the experiment moved on to the ITI. Stimuli were presented randomly until all items were shown.

Figure 1. Sample stimuli from the print, cursive, and assembled cursive conditions. (Please see text for details about assembled cursive.)

Results

Trials with voice-key errors or mispronunciations were removed from the data prior to analyses, constituting 7.1% (7.3% print; 6.9% cursive) and 6.5% (3.8% print; 9.0% cursive) of trials, respectively. RTs more than 2.5 *SD*s above and below the group means were also excluded from analyses, accounting for 4.2% of the correct RTs (3.1% print; 5.3% cursive). The resultant mean RTs are shown in Table 1; and the mean frequency and regularity effects (per condition) are shown in Figure 2. The results of item analyses are presented in Table 2.

Ex-Gaussian parameter estimates were derived for each participant in each cell of the design using the *egfit* MATLAB function written by Lacouture and Cousineau (2008). This function performs an iterative search of the RT distribution, reporting the three parameters (μ, σ, τ) of the probability distribution from which the observed RTs would most likely be sampled. Parameter estimates were first analyzed within each script condition using a series of 2 $\times$ 2 repeated-measures analyses of variance (ANOVAs) with within-subject factors frequency (high, low) and regularity (regular, irregular).

Beginning with the μ parameter in the computer print condition, we found a significant main effect of frequency, $F(1, 18) = 10.59$, $p < .01$, $\eta_p^2 = .37$, with an increased μ for low-frequency words. There was also a reliable regularity effect, $F(1, 18) = 5.98$, $p < .05$, $\eta_p^2 = .25$, with μ shifting higher for exception words. The Frequency $\times$ Regularity interaction was not reliable. Analyses of the σ and τ parameters yielded no reliable effects. In the cursive condition, the μ parameter produced only a reliable frequency effect, $F(1, 19) = 15.60$, $p = .001$, $\eta_p^2 = .45$, in the typical direction. The σ parameter also produced a reliable Frequency effect, $F(1, 19) = 5.58$, $p < .05$, $\eta_p^2 = .23$, with increased *SD*s for low-frequency words. Finally, τ produced a significant main effect

[2] The current research used cursive stimuli from only one writer. Future experiments should use a variety of stimuli in different hands to ensure that any observed effects are not the result of one individual's handwriting style. The "B" sections of each experiment partially address this issue by removing some of the "coarticulative" cues from the script.

Table 1
Experiments 1A and 1B, Mean Response Times, and Error Rates by Condition and Word Class

	Word class			
Condition	High-frequency regular	High-frequency exception	Low-frequency regular	Low-frequency exception
Computer print	529 (.06)	539 (.12)	559 (.09)	572 (.16)
Human cursive	659 (.11)	669 (.15)	722 (.18)	726 (.20)
Assembled cursive	660 (.15)	659 (.14)	709 (.19)	729 (.15)

of frequency, $F(1, 19) = 4.63$, $p < .05$, $\eta_p^2 = .20$, with greater skewing for low-frequency words.

Finally, we compared the distribution parameters in a series of $2 \times 2 \times 2$ mixed-model, repeated-measures ANOVAs with script as a between-subjects factor. For μ, we found a significant main effect of script, $F(1, 37) = 9.56$, $p < .01$, $\eta_p^2 = .21$, with the RT distribution shifting higher for cursive words. We also observed main effects of frequency, $F(1, 37) = 25.58$, $p < .001$, $\eta_p^2 = .41$, and regularity, $F(1, 37) = 5.19$, $p < .05$, $\eta_p^2 = .12$. In addition, there was a marginal three-way interaction of Frequency $\times$ Regularity $\times$ Script, $F(1, 37) = 3.51$, $p < .07$, $\eta_p^2 = .09$, with the traditional Frequency $\times$ Regularity interaction becoming more prominent in the cursive condition. Analyses of σ produced two reliable effects. The script effect was significant, $F(1, 37) = 5.72$, $p < .05$, $\eta_p^2 = .13$, with the cursive condition eliciting RT distributions with increased *SD*s. There was also a reliable Frequency effect, $F(1, 37) = 8.59$, $p < .01$, $\eta_p^2 = .19$, with increased *SD*s for low-frequency words. The τ parameter yielded similar results, with a main effect of script, $F(1, 37) = 34.14$, $p < .001$, $\eta_p^2 = .48$, because of increased distributional skewing for cursive words, and a reliable main effect of frequency, $F(1, 37) = 5.20$, $p < .05$, $\eta_p^2 = .12$.

Error Rates

In general, error rates (Table 1) showed the same patterns as μ in the distributional analyses. We conducted a repeated-measures ANOVA on error rates, with frequency and regularity as within-subject factors and script as a between-subjects factor. We observed a reliable effect of script, $F_S(1, 37) = 8.47$, $p < .01$, $\eta_p^2 = .19$, reflecting increased errors to the cursive words. A frequency effect was observed, $F_S(1, 37) = 29.24$, $p < .001$, $\eta_p^2 = .44$, as was a regularity effect, $F_S(1, 37) = 33.18$, $p < .001$, $\eta_p^2 = .47$, which interacted marginally with script, $F_S(1, 37) = 3.70$, $p < .07$, $\eta_p^2 = .09$, because of an increased regularity effect for printed words.

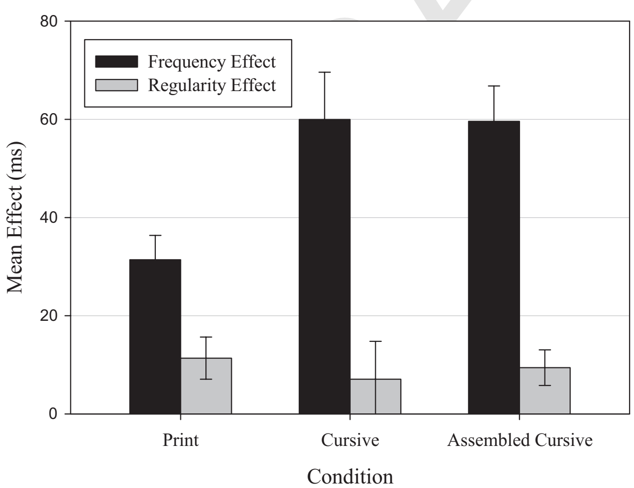

Figure 2. Mean effects of frequency and regularity as a function of script condition in Experiments 1A and 1B.

Discussion

The results of Experiment 1A suggest that handwriting slows down lexical access, allowing word frequency to more profoundly affect recognition. When people named handwritten words, the process was strongly affected by frequency, as previously reported by Manso De Zuniga et al. (1991). Although not evident in the distributional analyses, standard RT analyses produced a reliable Frequency $\times$ Script interaction, $F_S(1, 37) = 6.75$, $p < .05$, $\eta_p^2 = .15$, with an increased frequency effect in the cursive condition. This interaction did not emerge in the distributional analyses because of their partitioned nature. Balota and Spieler (1999) found that frequency effects reflected a combination of distributional shifting and skewing (manifested in the μ and τ parameters). As such, a reliable overall effect can become two separate null effects when divided for analysis. Nonetheless, the reliable effect in the overall analysis, coupled with the findings of Experiment 1B, clearly indicate stronger frequency effects for cursive words.[3] Fn3

In addition to the print and cursive conditions, we included a MiXeD CaSe condition in Experiment 1A to help ascertain whether cursive words add complications, beyond just perceptual

[3] In all our experiments, one potential concern was that, because RTs were considerably longer for handwritten words, any increased lexical effect to handwriting might reflect a mere scaling factor. For example, a frequency effect for handwriting might appear larger in absolute terms, but may be equivalent to the effect for print, in terms of proportions of their respective baselines. One method to address this is conduct standard ANOVAs on z-transformed data (Faust, Balota, Spieler, & Ferraro, 1999), setting all RTs to a common scale. We conducted such analyses for all the experiments in this report: Many of the reliable effects from the non-transformed ANOVAs remained, several became marginal. These results were easily understood in the context of the ex-Gaussian analyses. Specifically, those variables affecting shifting (μ) tended to remain in reliable after normalization, and those variables affecting skew (τ) tended to vanish. In this report, we focus on the ex-Gaussian analyses because they closely resembled the z-transformed ANOVAs while providing more meaningful results, allowing specific assessments of distribution shifting, in addition to changes in the upper tail.

Table 2
Experiments 1A and 1B, Item Analyses

Condition	Effect	Item statistics
Within-script analyses		
Print	**Frequency**	**$F(1, 156) = 37.21, p < .001, \eta_p^2 = .19$**
	Regularity	**$F(1, 156) = 6.21, p < .05, \eta_p^2 = .04$**
	Frequency × Regularity	$F(1, 156) = .45, p = .50, \eta_p^2 < .01$
Cursive	**Frequency**	**$F(1, 156) = 15.40, p < .001, \eta_p^2 = .09$**
	Regularity	$F(1, 156) = .10, p = .75, \eta_p^2 < .01$
	Frequency × Regularity	$F(1, 156) = .02, p = .89, \eta_p^2 < .01$
Assembled cursive	**Frequency**	$F(1, 156) = 18.37, p < .001, \eta_p^2 = .11$
	Regularity	$F(1, 156) = .43, p = .51, \eta_p^2 < .01$
	Frequency × Regularity	$F(1, 156) = .94, p = .34, \eta_p^2 < .01$
Between-script analyses		
Print and cursive	**Script**	**$F(1, 156) = 464.85, p < .001, \eta_p^2 = .75$**
	Frequency × Script	$F(1, 156) = 3.11, p = .08, \eta_p^2 = .02$
	Regularity × Script	$F(1, 156) = .46, p = .50, \eta_p^2 < .01$
	Frequency × Regularity × Script	$F(1, 156) = .02, p = .90, \eta_p^2 < .01$
Print and assembled cursive	**Script**	**$F(1, 156) = 554.12, p < .001, \eta_p^2 = .78$**
	Frequency × Script	**$F(1, 156) = 3.89, p = .05, \eta_p^2 = .02$**
	Regularity × Script	$F(1, 156) = .16, p = .69, \eta_p^2 < .01$
	Frequency × Regularity × Script	$F(1, 156) = .58, p = .45, \eta_p^2 < .01$

Note. Significant effects are reported in boldface.

novelty. If handwriting mainly disrupts the encoding of words' visual features, we should expect MiXeD CaSe words to create similar results. For example, Besner and Johnston (1989) reported increased frequency effects for MiXeD CaSe words. In our experiment, the MiXeD CaSe condition did not reliably differ from the standard print condition. Consequently, we refrain from further discussion of this additional condition, noting only that handwriting elicited response patterns that case-mixing did not.

In a separate, conceptually similar experiment, we also examined frequency and regularity effects for printed and cursive words using the stimuli from Seidenberg et al. (1984). Relative to Experiment 1A, the replication of Seidenberg et al. produced a larger increase in frequency effects, moving from print to cursive, and the effect was reflected largely in the μ parameter, with an effect size (η_p^2) increase of over 40%. In our replication of Seidenberg et al., although regularity effects were relatively small across stimulus forms (as was true in Experiment 1A), the well-known Frequency × Regularity interaction was selectively increased for handwritten cursive, in a standard analysis of cell means, $F_S(1, 38) = 9.43$, $p < .01$, $\eta_p^2 = .20$, but not in the ex-Gaussian analyses. According to the Dual-Route Cascaded model (DRC; Coltheart et al., 1993; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), lower-frequency words are processed via the "nonlexical route," which assembles pronunciation via grapheme-phoneme conversion rules. Regularity partly determines how quickly the word can be disambiguated. Low-frequency exception words will elicit competition between the lexical and non-lexical routes, delaying recognition. While these effects are often reported with standardized word forms, they increase when the stimuli are physically noisy, as in the case of handwriting. While Coltheart et al. (1993) did not speak directly to the effects of degraded or ambiguous inputs, one could infer that both pathways would be detrimentally affected by ambiguity to some extent.

Experiment 1B: Assembled Cursive

Taken at face value, one could easily assume that the effects elicited by handwritten stimuli arose because of differences in lexical processing. An alternative explanation could be that they arose because of differences in the production of the stimuli themselves. Inhoff (1991) reported that experienced typists were slower when typing low-frequency words, relative to high-frequency words. If the motor programs for generating key-presses are modulated by word frequency, so too could the motor programs for handwriting, leading to qualitative differences in the production of words from different lexical categories, and potentially affecting eventual reading fluency. Otherwise stated, it is possible that *production* fluency differed between low- and high-frequency words for our volunteer writer, potentially creating systematic variations in their perceptual quality. Because this was a concern for all of our experiments, we carried out companion ("B") experiments, designed to address this potential confound.

To eliminate the possibility that "authorship effects" differentially influenced lexical processing, we deconstructed the handwritten stimuli from Experiment 1A into an alphabet from which we could reconstruct each of the words. Essentially, we created a font using naturally handwritten items. By removing letters from their original contexts before reassembling them into words, any effects that were potentially built into the stimuli because of lexical factors were eliminated. These assembled cursive stimuli were used in a naming task identical to Experiment 1A.

Method

Subjects. Twenty Arizona State University students participated for partial course credit.

Stimuli. Using Adobe Photoshop, we decomposed the handwritten stimuli from Experiment 1A into a makeshift font. For each letter, three samples were taken: one from an instance where the letter appeared at the beginning of a word (e.g., first character), one from the middle, and one from the end (e.g., last character). Each sampled letter came from a different item. The stimuli from Experiment 1A were then re-generated using this "font." The sampled letters were used based on their locations within each item (e.g., letters sampled from the beginnings of words were always used at the beginnings of newly generated words). Connections between letters were added by hand to maintain the appearance of cursive handwriting (Figure 1).

Apparatus and procedure. The apparatus and procedure were identical to those of Experiment 1A.

Results

Trials with voice-key errors or mispronunciations were removed from the data prior to analyses, constituting 5.2% and 10.8% of trials, respectively. RTs greater than 2.5 *SD*s above and below the group mean were also excluded, accounting for 3.7% of trials. Table 1 shows the resultant cell means and Figure 2 shows the average Regularity and Frequency effects. Item analyses are reported in Table 2. Ex-Gaussian parameter estimates were calculated for each participant from each cell of the design. These parameter estimates were first analyzed in 2 × 2 repeated-measures ANOVAs with factors frequency and regularity. For the μ parameter, we found reliable main effects of frequency, $F(1, 19) = 5.07$, $p < .05$, $\eta_p^2 = .21$, and regularity, $F(1, 19) = 8.27$, $p = .01$, $\eta_p^2 = .30$, both in the typical directions. We also observed a marginal Frequency × Regularity interaction, $F(1, 19) = 3.84$, $p < .07$, $\eta_p^2 = .17$, with a larger regularity effect for low-frequency words. For the σ parameter, we found a marginal Regularity effect, $F(1, 19) = 4.33$, $p < .06$, $\eta_p^2 = .19$. The τ parameter revealed only a main effect of frequency, $F(1, 19) = 9.18$, $p < .01$, $\eta_p^2 = .33$, because of decreased distributional skewing for low-frequency words.

Finally, we compared the assembled cursive condition to the computer print condition from Experiment 1A using a 2 × 2 × 2 mixed-model repeated-measures ANOVA with between-subjects factor script. As there is great redundancy between these results and those from Experiment 1A, we report only the factors that interacted with script. Analysis of μ revealed a reliable main effect of script, $F(1, 37) = 17.56$, $p < .001$, $\eta_p^2 = .32$, with the distribution shifting higher for assembled cursive words. For σ, we found a marginal effect of script, $F(1, 37) = 3.97$, $p < .06$, $\eta_p^2 = .10$, with increased values for σ in the assembled cursive condition. The τ parameter produced a reliable script effect, $F(1, 37) = 32.32$, $p < .001$, $\eta_p^2 = .47$, with greater distributional skewing in the assembled cursive condition. There was also a reliable Frequency × Script interaction, $F(1, 37) = 4.25$, $p < .05$, $\eta_p^2 = .10$, with a larger frequency effect for assembled cursive words.

Discussion

The results of Experiment 1B corroborated those of Experiment 1A, replicating its results in a condition where potential authorship effects were eliminated. Taken together, both experiments demonstrate that handwritten words elicit larger top-down influences in recognition (as reflected by increased frequency effects), while having little effect on feed-forward processes (evidenced by null Regularity × Script interactions). In Experiment 2, we examined two variables that more explicitly reflect bottom-up and top-down influences in word perception: feed-forward and feedback consistency.

Experiment 2A: Bidirectional Consistency Effects in Lexical Decision

In Experiment 2, we investigated feed-forward and feedback consistency effects in lexical decision, extending Experiment 2 from Stone, Vanhoy, and Van Orden (1997) with a between-subjects contrast of word form, as in Experiment 1. The literature contains mixed findings regarding consistency effects on word perception: Most research has focused on feed-forward (henceforth FF) consistency, referring to the O-P translation of words. If multiple pronunciations accompany a word's spelling pattern, it is FF inconsistent (e.g., the orthographic body _INT, which is pronounced differently in PINT and MINT). If a spelling pattern can only be pronounced one way, the word is FF consistent. Consistent O-P mappings sometimes lead to faster RTs in perceptual tasks, however the effect varies from study to study and almost never emerges with high frequency words (Jared, McRae, & Seidenberg, 1990). Indeed, Plaut, McClelland, Seidenberg, and Patterson (1996) provided analytic proof that, in a connectionist model of word perception, frequency and consistency effects are naturally isomorphic to each other.

In a connectionist framework, consistency effects reflect differences across words in the statistical mapping of orthography to phonology. However, in such models, phonology must feedback activity to orthography: In theory, these reverse mappings should affect word perception, exactly like FF consistency. Stone et al. (1997) hypothesized that feedback (FB) consistency effects would originate from top-down processing of words: After orthographic patterns have been processed, feedback processes act as "fact-checkers" to verify that the generated pronunciations are consistent with their spellings. If a word's phonological body can be spelled in more than one way, the word is FB inconsistent (e.g., /_ip/ can be spelled _EEP or _EAP). If a *pronunciation* is only *spelled* one way, it is FB consistent. Note that consistency is not a binary variable, although it is often treated as one. Degrees of inconsistency vary according to the numbers and frequencies of neighboring words with shared or contrary pronunciations (friends and enemies; Jared et al., 1990).

Stone et al. (1997) compared lexical decision for fully consistent words and words that were only FB inconsistent, finding slower responses to the inconsistent words. In similar fashion, when all words were FB consistent, there was a significant FF consistency effect. Interestingly, they did not observe additive effects: RTs for words that were bidirectionally inconsistent did not differ from words that were only inconsistent in one direction. These findings were later replicated by Ziegler, Montant, and Jacobs (1997) and Lacruz and Folk (2004). In Experiment 2A, we used the same approach as Experiment 1, presenting the stimuli from Stone et al. (1997) in the same two formats. As before, handwriting was expected to complicate the initial visual processing of the words, thus increasing the need for top-down processing. In Experiment 1, frequency effects were increased for handwritten words. Following the analysis from Plaut et al. (1996), we expected FF consistency effects to be similarly increased in Experiment 2. And, by the logic of FB consistency (Stone et al., 1997), we should also expect

tapraid5/zfn-xhp/zfn-xhp/zfn00310/zfn2492d10z | xppws | S=1 | 3/20/10 | 11:05 | Art: 2008-1190

handwriting to increase FB consistency effects. That is, if feedback processes act as a spelling-checker (Van Orden, 1987), these processes should be less efficient when spelling patterns are "noisy," and thus harder to verify.

Method

Subjects. Eighty-six Arizona State University students participated for course credit (44 in the computer print condition, 42 in the human cursive condition).

Apparatus. Stimuli were presented via the E-prime program on Gateway computers with a screen resolution of 1,024 × 768. Responses were collected using a standard 5-button response box.

Stimuli. Stimuli were generated from the items used in Experiment 2 of Stone et al. (1997). These consisted of 40 FF consistent words and 40 FF inconsistent words. In each of these sets, half were FB consistent and half were FB inconsistent. All words were monosyllabic, and word lengths and frequencies were matched between cells. Eighty pronounceable nonwords (also from Stone et al.) were also used. They were created by replacing the initial consonant strings (heads) of monosyllabic words.

Procedure. Participants were randomly assigned to one of the script conditions and tested in groups up to eight people. Each participant was seated approximately 50 cm from the computer screen, and was instructed to quickly decide whether presented letter strings formed words or non-words, pressing the left and right buttons, respectively. The experiment began with a series of 10 practice trials to familiarize participants with the task and the word forms. Each trial began with a fixation point at the center of the screen for 750 ms, followed by the target probe. If no response was detected within 2.5 s, the words, "Too Late!" appeared on the screen for 500 ms, followed by a 1.5-s intertrial interval. After correct and incorrect responses, the words, "Correct!" and "Incorrect," appeared, respectively. Responses within 250 ms of the probe onset elicited the feedback, "TOO SOON!" Participants completed four blocks of 40 trials, separated by forced breaks. Each block contained 20 words (5 from each category) and 20 nonwords, presented randomly.

Results

Two participants were excluded from analyses (one from each script condition), because of having mean RTs greater than 2.5 *SD*s from their group means. Only correct responses were included in the analyses, excluding 13.6% of trials (9.6% print; 18% cursive). Correct trials with RTs over 2.5 *SD*s from the group means were also excluded, constituting 1.2% of trials (2.9% print; 2.5% cursive). The resultant mean RTs are shown in Table 3; mean forward and backward consistency effects (per condition) are shown in Figure 3. Item analyses are reported in Table 4. Ex-Gaussian parameter estimates were generated for each participant in each cell of the design, as before. First, we analyzed each script condition separately, using 2x2 repeated-measures ANOVAs with within-subject factors feed-forward and feedback consistency (denoted forward and backward, respectively). For μ in the print condition, there was a reliable backward effect, $F(1, 42) = 5.82$, $p < .05$, $\eta_p^2 = .12$. There was also a significant Forward × Backward interaction, $F(1, 42) = 5.83$, $p < .05$, $\eta_p^2 = .12$, with backward consistency influencing only forward consistent words. For σ there was only a reliable backward effect, $F(1, 42) = 5.21$, $p < .05$, $\eta_p^2 = .11$, with increased *SD* for backward inconsistent words. No reliable effects were observed for the τ parameter, suggesting that all significant effects were because of distributional shifting, not skewing. In the cursive condition, μ was the only parameter to yield any reliable effects. Within μ we found a reliable backward effect, $F(1, 40) = 8.00$, $p < .01$, $\eta_p^2 = .17$, with distributions shifting higher for backward inconsistent words.

Finally, we compared the print condition with the cursive condition, as before. For μ, we found a significant script effect, $F(1, 82) = 14.72$, $p < .001$, $\eta_p^2 = .15$, with the distribution shifting higher for cursive words. There was also a reliable backward effect, $F(1, 82) = 13.89$, $p < .001$, $\eta_p^2 = .15$, in the typical direction, and a reliable Forward × Backward interaction, $F(1, 82) = 4.29$, $p < .05$, $\eta_p^2 = .05$, with backward consistency only affecting forward consistent words. None of the interactions with script were reliable. Turning to σ, we found only a reliable main effect of backward consistency, $F(1, 82) = 4.76$, $p < .05$, $\eta_p^2 = .06$. For τ, there was a strong effect of script, $F(1, 82) = 43.18$, $p < .001$, $\eta_p^2 = .35$, because of increased distributional skewing for cursive words. Otherwise, no effects were reliable.

Error Rates

The error rates (Table 2) generally mirrored the μ distributional analyses, with the addition of a three-way interaction. We conducted a repeated-measures ANOVA on error rates, with forward and backward consistency as within-subject factors and script as a between-subjects factor. We observed a strong effect of script, $F_S(1, 82) = 46.13$, $p < .001$, $\eta_p^2 = .36$, with more errors to cursive words. We also observed a forward effect, $F_S(1, 82) = 52.86$, $p < .001$, $\eta_p^2 = .39$, and a reliable backward effect, $F_S(1, 82) = 74.52$, $p < .001$, $\eta_p^2 = .48$. The Forward × Backward interaction was also significant, $F_S(1, 82) = 35.10$, $p < .001$, $\eta_p^2 = .30$. Finally, the Forward × Backward × Script interaction was reliable, $F_s(1, 82) = 11.90$, $p = .001$, $\eta_p^2 = .13$, reflecting an increased interaction in the cursive condition.

Table 3
Experiments 2A and 2B, Mean Response Times and Error Rates by Condition and Word Class

	Word class			
Condition	FF consistent, FB consistent	FF consistent, FB inconsistent	FF inconsistent, FB consistent	FF inconsistent, FB inconsistent
Computer print	637 (.02)	694 (.13)	675 (.10)	672 (.17)
Human cursive	757 (.08)	825 (.23)	839 (.23)	822 (.22)
Assembled cursive	841 (.14)	912 (.26)	951 (.25)	935 (.30)

tapraid5/zfn-xhp/zfn-xhp/zfn00310/zfn2492d10z | xppws | S=1 | 3/20/10 | 11:05 | Art: 2008-1190

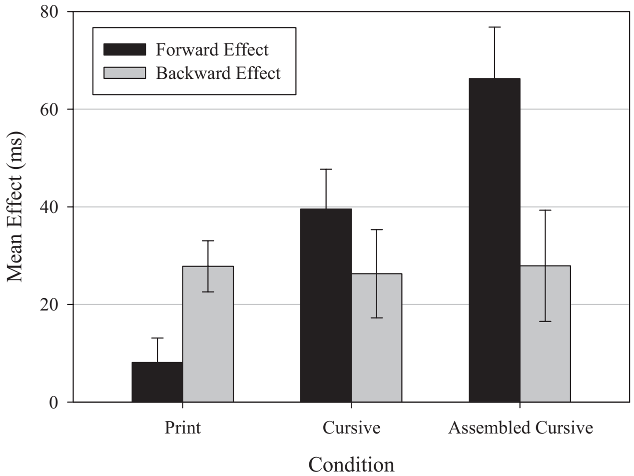

Figure 3. Mean effects of bidirectional consistency as a function of script condition in Experiments 2A and 2B.

Experiment 2B: Assembled Cursive

Method

Subjects. Forty-five Arizona State University students participated for course credit.

Stimuli, apparatus, and procedure. The stimuli from Experiment 2A were generated in an assembled cursive format, in the same fashion as Experiment 1B. The apparatus and procedure were identical to those from Experiment 2A.

Results

One participant was excluded from analysis for excessive errors. Only correct responses were included in the analyses, excluding 23% of the trials. Trials with RTs greater than 2.5 *SD*s above or below the group mean were also excluded, accounting for 2.8% of trials. See Table 3 for the resultant cell means and Figure 3 for the average forward and backward consistency effects. Item analyses are reported in Table 4. Ex-Gaussian parameter estimates were generated for each participant in each cell of the design, as before. For the μ parameter, there was a reliable main effect of forward consistency, $F(1, 43) = 15.16$, $p < .001$, $\eta_p^2 = .26$, with distributions shifted higher for inconsistent words. There was also a marginal Forward $\times$ Backward interaction, $F(1, 43) = 3.67$, $p < .07$, $\eta_p^2 = .08$. The σ parameter yielded only a significant forward effect, $F(1, 43) = 8.38$, $p < .01$, $\eta_p^2 = .16$, with increased *SD*s for forward inconsistent distributions. No reliable effects were observed for the τ parameter.

Finally, we compared the assembled cursive condition to the computer print condition from Experiment 2A. For μ, we found a reliable main effect of script, $F(1, 85) = 40.74$, $p < .001$, $\eta_p^2 = .32$, with distributions shifting higher for assembled cursive items. In addition, there was a Forward $\times$ Script interaction, $F(1, 85) = 13.95$, $p < .001$, $\eta_p^2 = .14$, because of an increased forward effect in the assembled cursive condition. Analysis of σ revealed a reliable main effect of script, $F(1, 85) = 20.99$, $p < .001$, $\eta_p^2 = .20$, with σ estimates increasing for assembled cursive items. We also found a significant forward effect, $F(1, 85) = 6.94$, $p = .01$, $\eta_p^2 = .08$, with increased *SD*s for forward inconsistent distributions, and this effect interacted reliably with script, $F(1, 85) = 5.67$, $p < .05$, $\eta_p^2 = .06$, because of a magnified forward effect in the assembled cursive condition. For τ, we found only a reliable main effect of script, $F(1, 85) = 49.31$, $p < .001$, $\eta_p^2 = .37$, with increased distributional skewing for assembled cursive items.

Error Rates

We compared error rates (Table 3) between the assembled cursive condition in Experiment 2B and the computer print con-

Table 4
Experiments 2A and 2B, Item Analyses

Condition	Effect	Item statistics
Within-script analyses		
Print	Forward	$F(1, 76) = 1.39$, $p = .24$, $\eta_p^2 = .02$
	Backward	$\mathbf{F(1, 76) = 6.12, p < .05, \eta_p^2 = .08}$
	Forward × Backward	$\mathbf{F(1, 76) = 4.17, p < .05, \eta_p^2 = .05}$
Cursive	Forward	$F(1, 76) = 1.56$, $p = .22$, $\eta_p^2 = .02$
	Backward	$F(1, 76) = 2.32$, $p = .13$, $\eta_p^2 = .03$
	Forward × Backward	$\mathbf{F(1, 76) = 5.28, p < .05, \eta_p^2 = .07}$
Assembled cursive	**Forward**	$F(1, 76) = 7.35$, $p < .01$, $\eta_p^2 = .08$
	Backward	$F(1, 76) = .90$, $p = .35$, $\eta_p^2 = .01$
	Forward × Backward	$F(1, 76) = 2.77$, $p = .10$, $\eta_p^2 = .04$
Between-script analyses		
Print and cursive	**Script**	$\mathbf{F(1, 76) = 182.01, p < .001, \eta_p^2 = .71}$
	Forward × Script	$F(1, 76) = .49$, $p = .49$, $\eta_p^2 = .01$
	Backward × Script	$F(1, 76) = .05$, $p = .83$, $\eta_p^2 < .01$
	Forward × Backward × Script	$F(1, 76) = 1.86$, $p = .18$, $\eta_p^2 = .02$
Print and assembled cursive	**Script**	$F(1, 76) = 533.67$, $p < .001$, $\eta_p^2 = .88$
	Forward × Script	$F(1, 76) = 5.29$, $p < .05$, $\eta_p^2 = .07$
	Backward × Script	$F(1, 76) = .30$, $p = .58$, $\eta_p^2 < .01$
	Forward × Backward × Script	$F(1, 76) = .30$, $p = .59$, $\eta_p^2 < .01$

Note. Significant effects are reported in boldface.

dition in Experiment 2A. We observed a reliable effect of script, $F_S(1, 80) = 33.38$, $p < .001$, $\eta_p^2 = .29$, with more errors to assembled cursive words. We also observed a reliable forward effect, $F_S(1, 80) = 61.37$, $p < .001$, $\eta_p^2 = .43$; a backward effect, $F_S(1, 80) = 76.84$, $p < .001$, $\eta_p^2 = .49$; and a Forward $\times$ Backward interaction, $F_S(1, 80) = 7.35$, $p < .01$, $\eta_p^2 = .08$. None of these effects interacted reliably with script.

Discussion

Taken together, the results from Experiments 2A and 2B showed clear effects of stimulus format, but did not entirely conform to expectations. Given noisier visual input, we expected increased feedback consistency effects, as top-down "spelling verification" would be slower with less pristine characters. Instead, the effect was fairly stable across conditions. Feed-forward consistency effects, however, were weak in the computer print condition, growing more robust in the processing of human cursive (in the assembled cursive condition). We consider these results further after Experiment 3. One concern with using lexical decision is that word RTs can be unduly affected by the difficulty of nonword rejection. Indeed, a subsequent analysis of Experiment 2A revealed an inflated false-alarm rate (i.e., calling nonwords "words") in the human cursive condition, $F(1, 83) = 44.40$, $p < .001$, $\eta_p^2 = .35$, complicating direct comparisons across word form conditions. Thus, Experiments 3 and 4 both employed a naming paradigm.

Experiment 3A: Bidirectional Consistency Effects in Naming

In an attempt to further ascertain the influence of handwritten word forms on consistency effects, we again used the word stimuli from Stone et al. (1997), now in a naming task. Because FB consistency effects have received little attention overall, few studies have assessed their influence in naming. An exception is Lacruz and Folk (2004) who, contrary to Stone et al., found significant FB effects for both high- and low-frequency words. Our expectations followed those for Experiment 2: If forward consistency and word frequency are coupled, as Plaut et al. (1996) assert, forward effects should be magnified in naming of cursive words in the same way that frequency effects were magnified in Experiment 1. We also predicted that FB consistency effects would increase when stimuli were physically noisy, reflecting the relative difficulty of spelling verification.

Method

Subjects. Forty Arizona State University students participated for course credit (20 in the print condition, 20 in the cursive condition). All were native English speakers, and none had participated previously.

Stimuli, apparatus, and procedure. The stimuli (in computer print and human cursive forms) were identical to those used in Experiment 2A, excluding the nonwords. The apparatus and procedure were identical to those of Experiment 1.

Results and Discussion

Three participants were excluded from analysis (two from the print condition and one from the cursive condition) because of having average RTs greater than 2.5 *SD*s above their group means or excessive error rates. Any trials containing mispronunciations or voice-key errors were excluded from analysis, comprising 8.5% (5.8% print; 11.1% cursive) and 3.9% (2.6% print; 5.1% cursive) of trials, respectively. Trials with RTs greater than 2.5 *SD*s from the group means were also excluded, including 2.9% of the correct trials (2% print; 3.8% cursive). The resultant mean RTs are shown in Table 5; mean forward and backward consistency effects (per T5
condition) are shown in Figure 4. Item analyses are reported in F4
Table 6. Ex-Gaussian parameter estimates were generated for each T6
participant in each cell of the design. Our analyses began with a series of 2 $\times$ 2 repeated-measures ANOVAs (one per parameter, script condition) with within-subject factors forward and backward. In the computer print condition, the μ parameter produced only a reliable main effect of backward consistency, $F(1, 17) = 4.78$, $p < .05$, $\eta_p^2 = .22$, because of higher distributional shifting for backward inconsistent words. No reliable effects were observed for the σ or τ parameters. Turning to μ in the cursive condition, we found a reliable forward effect, $F(1, 18) = 5.46$, $p < .05$, $\eta_p^2 = .23$, with distributions shifting higher for forward inconsistent words. The σ parameter produced a significant Forward $\times$ Backward interaction, $F(1, 18) = 7.33$, $p < .05$, $\eta_p^2 = .29$, with inconsistency in either direction leading to greater *SD*s. The τ parameter also produced a reliable Forward $\times$ Backward interaction, $F(1, 18) = 17.54$, $p = .001$, $\eta_p^2 = .49$, with inconsistency in either direction leading to increased distributional skewing.

We next compared the print condition to the cursive condition on all three parameters. Analysis of μ produced a main effect of script, $F(1, 35) = 11.66$, $p < .01$, $\eta_p^2 = .25$, with larger μ for cursive words. There was a reliable backward effect, $F(1, 35) = 4.49$, $p < .05$, $\eta_p^2 = .11$, because of higher distributional shifting for backward inconsistent words, and there was a reliable Forward $\times$ Script interaction, with forward effects increasing from the print to the cursive condition. Analysis of σ produced a reliable Forward $\times$ Backward interaction, $F(1, 35) = 6.61$, $p < .05$, $\eta_p^2 = .16$, which also increased from the print to cursive condition, as evidenced by a reliable three-way Forward $\times$ Backward $\times$ Script

Table 5
Experiments 3A and 3B, Mean Response Times and Error Rates by Condition and Word Class

	Word class			
Condition	FF consistent, FB consistent	FF consistent, FB inconsistent	FF inconsistent, FB consistent	FF inconsistent, FB inconsistent
Computer print	545 (.01)	571 (.10)	559 (.08)	582 (.15)
Human cursive	679 (.06)	748 (.17)	739 (.17)	759 (.25)
Assembled cursive	695 (.06)	780 (.17)	743 (.09)	793 (.28)

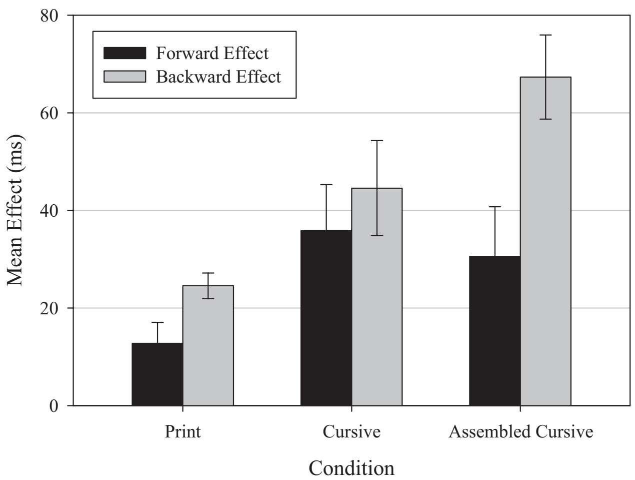

Figure 4. Mean effects of bidirectional consistency as a function of script condition in Experiments 3A and 3B.

interaction, $F(1, 35) = 5.23$, $p < .05$, $\eta_p^2 = .13$. The τ parameter produced a reliable script effect, $F(1, 35) = 48.47$, $p < .001$, $\eta_p^2 = .58$, with increased distributional skewing in the cursive condition. In addition, the Forward × Backward interaction was reliable, $F(1, 35) = 11.35$, $p < .01$, $\eta_p^2 = .25$, and the size of the interaction was dependent on script, as reflected by a significant Forward × Backward × Script interaction, $F(1, 35) = 48.47$, $p < .001$, $\eta_p^2 = .58$.

Error Rates

A repeated-measures ANOVA on error rates, with forward and backward consistency as within-subject factors and script as a between-subjects factor, produced a reliable effect of script, $F_s(1, 35) = 11.42$, $p < .01$, $\eta_p^2 = .25$, with more errors in the cursive condition. We observed a reliable forward effect, $F_s(1, 35) = 52.97$, $p < .001$, $\eta_p^2 = .60$, and a reliable backward effect, $F_s(1, 35) = 53.99$, $p < .001$, $\eta_p^2 = .61$. The Forward × Backward interaction was not reliable, and neither interacted with script.

Experiment 3B: Assembled Cursive

Method

Subjects. Twenty-seven Arizona State University students participated for course credit.

Stimuli, apparatus, and procedure. The stimuli were identical to those used in Experiment 2B, excluding the nonwords. The apparatus and procedure were identical to those from Experiment 1A.

Results

Two participants were excluded from analyses because of having RTs greater than 2.5 *SD*s above the group average or excessive error rates. Only accurate trials were analyzed, excluding 11% of trials because of mispronunciations, and 4% because of voice key errors. RTs greater than 2.5 *SD*s from the group mean were also excluded, constituting 4% of correct trials. The resultant mean RTs are shown in Table 5; mean forward and backward consistency effects are shown in Figure 4. Item analyses are reported in Table 6. In keeping with the previous experiments, parameter estimates were obtained for each participant in each cell of the design. We first analyzed each parameter separately via 2 × 2 repeated-measures ANOVAs with within-subject factors forward and backward. The μ parameter produced a reliable backward effect, $F(1, 24) = 6.62$, $p < .05$, $\eta_p^2 = .22$, with distributions shifting higher for backward inconsistent words. There was also a reliable Forward × Backward interaction, $F(1, 24) = 11.37$, $p < .01$, $\eta_p^2 = .32$, with increased values of μ for inconsistency in either direction. The σ parameter yielded no reliable effects, but analysis of τ

Table 6
Experiments 3A and 3B, Item Analyses

Condition	Effect	Item statistics
Within-script analyses		
Print	Forward	$F(1, 76) = 1.53, p = .22, \eta_p^2 = .02$
	Backward	$\mathbf{F(1, 76) = 7.58, p < .01, \eta_p^2 = .09}$
	Forward × Backward	$F(1, 76) = .73, p = .40, \eta_p^2 = .01$
Cursive	Forward	$F(1, 76) = 2.72, p = .10, \eta_p^2 = .04$
	Backward	$\mathbf{F(1, 76) = 4.76, p < .05, \eta_p^2 = .06}$
	Forward × Backward	$F(1, 76) = .78, p = .38, \eta_p^2 = .01$
Assembled cursive	Forward	$F(1, 76) = 1.88, p = .18, \eta_p^2 = .02$
	Backward	$\mathbf{F(1, 76) = 10.94, p = .001, \eta_p^2 = .13}$
	Forward × Backward	$F(1, 76) = .76, p = .39, \eta_p^2 = .01$
Between-script analyses		
Print and cursive	**Script**	$\mathbf{F(1, 76) = 181.99, p < .001, \eta_p^2 = .71}$
	Forward × Script	$F(1, 76) = 2.07, p = .15, \eta_p^2 = .03$
	Backward × Script	$F(1, 76) = 2.41, p = .13, \eta_p^2 = .03$
	Forward × Backward × Script	$F(1, 76) = .51, p = .48, \eta_p^2 = .01$
Print and assembled cursive	**Script**	$\mathbf{F(1, 76) = 443.90, p < .001, \eta_p^2 = .85}$
	Forward × Script	$F(1, 76) = .81, p = .37, \eta_p^2 = .01$
	Backward × Script	$\mathbf{F(1, 76) = 5.20, p < .05, \eta_p^2 = .06}$
	Forward × Backward × Script	$F(1, 76) = .29, p = .59, \eta_p^2 < .01$

Note. Significant effects are reported in boldface.

produced a marginal Forward × Backward interaction, $F(1, 24) = 3.75$, $p < .07$, $\eta_p^2 = .14$.

We also compared the assembled cursive condition to the computer print condition from Experiment 3A. Examination of the μ parameter produced a reliable script effect, $F(1, 41) = 13.28$, $p = .001$, $\eta_p^2 = .25$, with μ increasing in the assembled cursive condition. The Forward × Script interaction was null, but the Forward × Backward interaction was reliable, $F(1, 41) = 6.07$, $p < .05$, $\eta_p^2 = .13$. This interaction increased from the print to assembled cursive condition, as evidenced by a reliable three-way Forward × Backward × Script interaction, $F(1, 41) = 4.42$, $p < .05$, $\eta_p^2 = .10$. For the σ parameter, we observed only a main effect of script, $F(1, 41) = 7.68$, $p < .01$, $\eta_p^2 = .16$, with increased *SD*s for assembled cursive words. The same was true for the τ parameter; only a main effect of script was reliable, $F(1, 41) = 33.02$, $p < .001$, $\eta_p^2 = .45$, because of increased skewing for assembled cursive words.

Error Rates

A comparison of error rates (Table 5) between the computer print and assembled cursive conditions revealed a forward effect, $F_S(1, 49) = 58.38$, $p < .001$, $\eta_p^2 = .54$; a backward effect, $F_S(1, 49) = 75.56$, $p < .001$, $\eta_p^2 = .61$; and a Forward × Backward interaction, $F_S(1, 49) = 6.72$, $p < .05$, $\eta_p^2 = .12$. We also observed a reliable Backward × Script interaction, because of increased backward effects in the assembled cursive condition.

Discussion

Taken together, the results of Experiments 2 and 3 provide a clearer picture of FF and FB consistency effects, suggesting that they are task-dependent. In lexical decision, which does not necessarily require full access to phonology, we observed a large increase in forward effects, moving from computer print to human cursive. This was not true of FB consistency: Although the effect was observed in both script conditions, it did not differ with the addition of physical ambiguity. When we required participants to access phonology completely by using a naming task, not only did we observe larger FF effects, but FB effects were also systematically increased. A possible explanation for the discrepancy across Experiments 2 and 3 is provided by Balota et al. (2004), who conducted a large regression study, examining the myriad factors affecting word recognition. They assessed the influence of many lexical variables, including forward and backward consistency, in both naming and lexical decision. Balota et al. found that bidirectional consistency effects were substantially decreased in lexical decision, suggesting that people are less sensitive to phonological manipulations when they are only required to make
AQ: 2 word/nonword judgments. Balota and Chumbley (1984) suggested that, when making lexical decisions, people can assess overall familiarity, reducing the potential impact of variables that modulate access to phonological codes.

In a recent analysis of FB consistency effects, Ziegler, Petrova, and Ferrand (2008) concluded, after a series of null findings, that, "there is very little evidence, neither empirically nor computationally, for feedback consistency effects in the visual modality," (p. 17). Certainly, the literature on feedback effects is contradictory (Massaro & Jesse, 2005; Peereman, Content, & Bonin, 1998; Ziegler et al., 2008), but no previous researchers have examined these effects under conditions of orthographic ambiguity. Ziegler et al. noted that the use of normalized visual stimuli allows for easy "clean-up" of the input in an interactive model. This sort of clean-up (i.e., backwards matching) is impossible in the auditory modality, wherein they found strong feedback consistency effects. It follows from this argument that the natural ambiguity of handwritten stimuli should make such clean-up more difficult, allowing feedback to influence recognition. This is exactly what we found, again suggesting that handwritten word recognition may be considered an analogue to spoken word recognition.

In addition to increased FB consistency effects, we also observed increased FF consistency effects in the human cursive condition. At first blush, this increase seems surprising, because FF consistency is most naturally classified as a "bottom-up" variable, affecting the translation of print to sound. However, in a formal analysis of their connectionist model of reading, Plaut et al. (1996) showed that frequency and consistency converge on a common process, with increases in either driving processing in the network toward its asymptote. Frequency and consistency also trade off with each other, such that increases in either parameter can mitigate low values in the other. Thus, although FF consistency is a "bottom-up" variable, its behavioral effects are partly determined by variations in "top-down" word frequency, and vice versa. In Experiment 1, we found that frequency effects were magnified among handwritten words. In Experiment 3, frequency was held at uniform low values, potentially allowing consistency effects to more clearly emerge. By presenting words in handwritten form, we presumably slowed the translation of print to sound, allowing modest consistency effects (with computer print) to become more robust.

Experiment 4A: Semantic Effects

Experiments 1 through 3 all included variables thought to index top-down processing. We consistently observed magnified effects of top-down lexical variables with handwritten stimuli. Our final two experiments focused on imageability, a variable that is undeniably top-down in nature. Connectionist models of word recognition allow for the top-down influence of meaning on activation of phonology (Strain & Herdman, 1999), although Strain, Patterson, and Seidenberg (1995) described semantic processing as a "last-ditch" effort in word naming. In the well-known triangle model (Seidenberg & McClelland, 1989), word naming can be viewed as a set of cascaded processes, with semantics occasionally having an impact. Presentation of a letter string activates phonology, but also activates semantic representations. In most cases, naming is possible based on rapidly-forming resonance between print and sound, with little opportunity for semantics to affect performance. If, however, mapping from print to sound is slow (e.g., because of FF inconsistency), top-down influences from semantics may emerge.

Strain et al. (1995) provided evidence for a complex role of semantics in word naming. In their experiment, people named words that were either low-frequency regular or low-frequency exceptions. Different word sets also varied in imageability, defined as the extent to which words elicit mental images (Toglia & Battig, 1978). This is a commonly used semantic variable, because of its utility as a predictor of reading ability in people with deep dyslexia

tapraid5/zfn-xhp/zfn-xhp/zfn00310/zfn2492d10z | xppws | S=1 | 3/20/10 | 11:05 | Art: 2008-1190

(Allport & Funnell, 1981). In Strain et al. (1995), word regularity and imageability interacted, with slowest naming RTs to exception words with abstract (i.e., less imageable) meanings. These semantic effects were replicated by Shibahara, Zorzi, Hill, Wydell, and Butterworth (2003). Harm and Seidenberg (2004) presented stimuli similar to those used by Strain et al. (1995) to an elaborated triangle model of lexical access. The model replicated the results of Strain et al., with increased imageability effects for low-frequency exception words. In a similar study, Strain and Herdman (1999) found that semantic effects were larger for participants with lower phonological reading abilities. Among such readers, semantic effects extend to regular words, much as they do in cases of extreme phonological impairment, such as deep dyslexia. In similar fashion, handwriting may also force greater reliance on semantic feedback in naming (Van Orden & Goldinger, 1994).

Monaghan and Ellis (2002) took issue with much of the previous research on semantic effects in word naming, arguing that none of the designs controlled for *age of acquisition* (AoA) of the stimulus item. Words that are learned earlier in childhood reliably elicit faster RTs in naming tasks, and many of the earliest learned words are high-imageability words. Monaghan and Ellis reported that, in a reanalysis of the Strain et al. (1995) data with AoA as a covariate, the effect of imageability disappeared. In addition, when the words are balanced for AoA, no imageability effect occurs. Shibahara et al. (2003) replicated Experiment 1 of Strain et al. (1995), also finding that semantic effects disappeared when AoA was included as a covariate. But when they used a new, larger word list containing only low-frequency words, the imageability–regularity interaction emerged once again, suggesting an influence of imageability beyond AoA. More recently, Cortese and Khanna (2007) included AoA ratings in a mega-study of both naming and lexical decision times, observing that AoA accounts for substantially more variance than a predictor set of common lexical variables. Notably, imageability no longer predicted naming times after AoA effects had been partialled out (although it still accounted for unique variance in lexical decision).

In Experiment 4, we replicated Experiment 2 of Strain et al. (1995) with the addition of varying word forms, as in the previous experiments. From the perspective of Harm and Seidenberg's (2004) model, ambiguity in word forms is expected to magnify the influence of semantics: Handwriting should slow the build-up of activation between orthography and phonology, encouraging a reliance on semantic feedback to help resolve the input. In addition, we collected AoA ratings for the stimulus set from Strain et al., to include as covariates in item analyses, allowing us to examine whether imageability produces an effect beyond AoA.

Method

Subjects. Forty Arizona State University students participated in the naming task (20 in the print condition, 20 in the cursive condition) and 47 additional students provided AoA ratings in exchange for course credit.

Stimuli. The pool of 64 stimuli from Strain et al. (1995) consisted entirely of low-frequency words, with a 2×2 design of Regularity (irregular, regular) by Imageability (low, high). Words were matched in quartets for frequency, number of letters, and class of initial phoneme. Stimuli in the two script styles were generated in the same fashion as those for Experiments 1 through 3.

Apparatus and procedure. The apparatus and procedure for the naming task were identical to those of Experiments 1 and 3. The 64 stimuli were presented randomly in two blocks of 32 trials. The protocol used by Cortese and Khanna (2007, 2008) was employed in the AoA ratings task. Ratings were collected via the E-prime program on Gateway computers with a screen resolution of $1{,}024 \times 768$. Responses were recorded using a standard keyboard number pad.

Results

All trials with voice-key errors or mispronunciations were excluded from analysis, accounting for 3.2% (2.6% print; 3.9% cursive) and 11.9% (10.2% print; 13.7% cursive) of trials. Trials with RTs greater than 2.5 *SD*s from the group means were also excluded, constituting 2.8% (2.4% print; 3.2% cursive) of accurate trials. The resultant mean RTs are shown in Table 7; mean regularity and imageability effects (per condition) are shown in Figure 5. Item analyses are reported in Table 8. As in the previous experiments, ex-Gaussian parameter estimates were generated for each participant in each cell of the design. We first analyzed each script condition separately with a series of 2×2 repeated-measures ANOVAs (one per parameter) with within-subject factors Regularity (regular, irregular) and Imageability (high, low). In the print condition, the μ parameter produced a reliable Regularity effect, $F(1, 19) = 5.36$, $p < .05$, $\eta_p^2 = .22$, with higher μ values for irregular words. There was also an Imageability effect, $F(1, 19) = 12.22$, $p < .01$, $\eta_p^2 = .39$, with distributions shifted higher for low-imageability words. In addition, there was a marginal Regularity $\times$ Imageability interaction, $F(1, 19) = 4.33$, $p < .06$, $\eta_p^2 = .19$, in the typical direction, with a larger imageability effect for irregular words. Neither the σ nor the τ parameters produced any reliable effects.

T7 F5, T8

Table 7
Experiments 4A and 4B, Mean Response Times and Error Rates By Condition and Word Class

	Word class			
Condition	Regular high-imageability	Regular low-imageability	Exception high-imageability	Exception low-imageability
Computer print	586 (.03)	600 (.07)	590 (.11)	644 (.29)
Human cursive	738 (.08)	773 (.17)	728 (.17)	873 (.29)
Assembled cursive	765 (.09)	859 (.18)	782 (.21)	901 (.40)

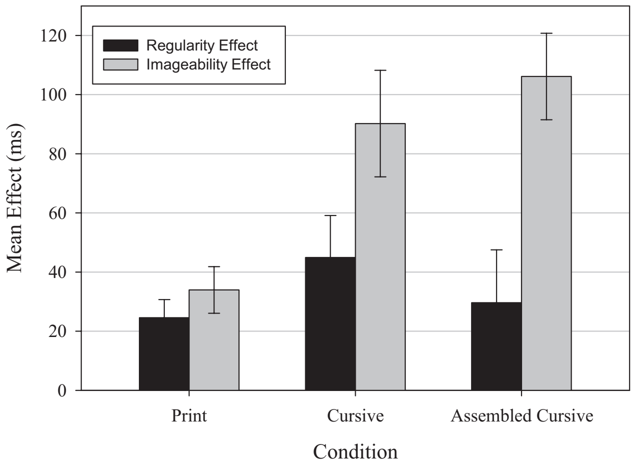

Figure 5. Mean effects of regularity and imageability as a function of script condition in Experiments 4A and 4B.

In the cursive condition, the μ parameter produced three significant effects. The Regularity effect was reliable, $F(1, 19) = 4.49$, $p < .05$, $\eta_p^2 = .19$, with higher μ values for irregular words. The imageability effect was robust, $F(1, 19) = 27.56$, $p < .001$, $\eta_p^2 = .59$, with the distribution shifted higher for low-imageability words. The Regularity × Imageability interaction was also reliable, $F(1, 19) = 5.83$, $p < .05$, $\eta_p^2 = .24$, because of increased imageability effects for irregular words. Analysis of σ produced only an Imageability effect, $F(1, 19) = 4.47$, $p < .05$, $\eta_p^2 = .19$, because of increased *SD*s for low-imageability words. The τ parameter produced no reliable effects.

Comparing the print condition to the cursive condition, for μ, there was a significant script effect, $F(1, 38) = 8.67$, $p < .01$, $\eta_p^2 = .19$, with increased values of μ in the cursive condition. The Regularity effect was reliable, $F(1, 38) = 8.92$, $p < .01$, $\eta_p^2 = .19$, as was the Imageability effect, $F(1, 38) = 39.36$, $p < .001$, $\eta_p^2 = .52$ (both in the typical directions), and Regularity interacted with Imageability, $F(1, 38) = 9.52$, $p < .01$, $\eta_p^2 = .20$. The imageability effect increased from the print to the cursive condition, as evidenced by a Imageability × Script interaction, $F(1, 38) = 9.09$, $p < .01$, $\eta_p^2 = .19$. The σ parameter produced only an imageability effect, $F(1, 38) = 5.05$, $p < .05$, $\eta_p^2 = .12$. For τ, the only significant finding was a script effect, $F(1, 38) = 44.41$, $p < .001$, $\eta_p^2 = .54$, because of increased skewing in the cursive condition.

Error Rates

We conducted a repeated-measures ANOVA on error rates (Table 7), with Regularity and Imageability as within-subject factors and script as a between-subjects factor. Error rates varied marginally between the script conditions, with the cursive condition producing more errors, $F_S(1, 38) = 3.83$, $p < .06$, $\eta_p^2 = .09$. We observed a Regularity effect, $F_S(1, 38) = 108.40$, $p < .001$, $\eta_p^2 = .74$, and an Imageability effect, $F_S(1, 38) = 56.03$, $p < .001$, $\eta_p^2 = .60$. Finally, the Regularity × Imageability interaction was significant, $F_S(1, 38) = 10.25$, $p < .01$, $\eta_p^2 = .21$, and it decreased slightly in the cursive condition, as reflected by a Script × Regularity × Imageability interaction, $F_S(1, 38) = 3.78$, $p < .06$, $\eta_p^2 = .09$.

AoA Ratings

Per Cortese and Khanna (2008), AoA ratings with RTs below 500 ms were excluded from analyses. Average AoA ratings for

Table 8
Experiments 4A and 4B, Item Analyses

Condition	Effect	Item statistics
Within-script analyses		
Print	**Age of acquisition**	$\mathbf{F(1, 59) = 12.67, p = .001, \eta_p^2 = .18}$
	Regularity	$\mathbf{F(1, 59) = 9.06, p < .01, \eta_p^2 = .13}$
	Imageability	$F(1, 59) < .01, p = .97, \eta_p^2 < .01$
	Regularity × Imageability	$\mathbf{F(1, 59) = 4.23, p < .05, \eta_p^2 = .07}$
Cursive	**Age of acquisition**	$\mathbf{F(1, 59) = 12.10, p = .001, \eta_p^2 = .17}$
	Regularity	$F(1, 59) = 3.59, p = .06, \eta_p^2 = .06$
	Imageability	$F(1, 59) = .05, p = .83, \eta_p^2 < .01$
	Regularity × Imageability	$F(1, 59) = 3.55, p = .07, \eta_p^2 = .06$
Assembled cursive	**Age of acquisition**	$\mathbf{F(1, 59) = 6.85, p < .05, \eta_p^2 = .10}$
	Regularity	$F(1, 59) = 2.28, p = .14, \eta_p^2 = .04$
	Imageability	$F(1, 59) = .30, p = .59, \eta_p^2 < .01$
	Regularity × Imageability	$F(1, 59) = .05, p = .83, \eta_p^2 < .01$
Between-script analyses		
Print and cursive	Script	$F(1, 59) = .83, p = .37, \eta_p^2 = .01$
	Age of acquisition × Script	$\mathbf{F(1, 59) = 3.85, p = .05, \eta_p^2 = .06}$
	Regularity × Script	$F(1, 59) = .34, p = .56, \eta_p^2 < .01$
	Imageability × Script	$F(1, 59) = .06, p = .81, \eta_p^2 < .01$
	Regularity × Imageability × Script	$F(1, 59) = 1.01, p = .32, \eta_p^2 = .02$
Print and assembled cursive	Script	$F(1, 59) = 2.70, p = .11, \eta_p^2 = .04$
	Age of acquisition × Script	$F(1, 59) = 2.26, p = .14, \eta_p^2 = .04$
	Regularity × Script	$F(1, 59) = .25, p = .62, \eta_p^2 < .01$
	Imageability × Script	$F(1, 59) = .38, p = .54, \eta_p^2 = .01$
	Regularity × Imageability × Script	$F(1, 59) = .30, p = .59, \eta_p^2 = .01$

Note. Significant effects are reported in boldface.

each word were entered as a covariate in mixed-model item ANOVAs, with script as a within-subject factor and Regularity and Imageability as between-subjects factors. The results of the item analyses for each script and the interactions between scripts can be seen in Table 8. The omnibus ANOVA produced a reliable AoA effect, $F_I(1, 59) = 17.68$, $p < .001$, $\eta_p^2 = .23$, as well as a reliable Regularity effect, $F_I(1, 59) = 7.20$, $p < .01$, $\eta_p^2 = .11$. Although removing shared variance with AoA negated the main effect of imageability, the Regularity × Imageability interaction was still reliable, $F_I(1, 59) = 5.40$, $p < .05$, $\eta_p^2 = .08$, suggesting an effect of imageability beyond AoA. The same interaction approached significance, $F(1, 59) = 3.55$, $p = .07$, $\eta_p^2 = .06$, in the human cursive condition.

Experiment 4B: Assembled Cursive

Method

Subjects. Twenty-seven Arizona State University students participated for course credit.

Stimuli, apparatus, and procedure. The stimuli from Experiment 4A were generated in an assembled cursive format, as in prior experiments. The apparatus and procedure were identical to those from Experiment 4A.

Results

Trials with voice-key errors or mispronunciations were removed from the data prior to analysis, constituting 3.8% and 18.3% of trials, respectively. Trials with RTs greater than 2.5 *SD*s from the group mean were also excluded, accounting for 3.8% of accurate trials. The resultant mean RTs are shown in Table 7; mean regularity and imageability effects are shown in Figure 5. Item analyses are reported in Table 8. In keeping with previous experiments, ex-Gaussian parameter estimates were generated for each participant in each cell of the design. We analyzed the assembled cursive items using 2 × 2 repeated measures ANOVAs (one per parameter). For μ, there was a robust Imageability effect, $F(1, 26) = 34.20$, $p < .001$, $\eta_p^2 = .57$, in the typical direction. There was also a reliable Regularity × Imageability interaction, $F(1, 26) = 4.38$, $p < .05$, $\eta_p^2 = .14$, with a larger imageability effect for irregular words. The σ and τ parameters produced no significant effects.

Finally, we compared the assembled cursive condition to the print condition from Experiment 4A. The μ parameter produced a reliable main effect of script, $F(1, 45) = 13.86$, $p = .001$, $\eta_p^2 = .24$, with increased μ values in the assembled cursive condition. As in the previous experiment, Imageability interacted with script, $F(1, 45) = 11.99$, $p = .001$, $\eta_p^2 = .21$, with a larger imageability effect in the assembled cursive condition. The σ parameter revealed only a main effect of script, $F(1, 45) = 5.82$, $p < .05$, $\eta_p^2 = .12$, with σ increasing in the assembled cursive condition. The same was true of the τ parameter, with only a script effect, $F(1, 45) = 39.94$, $p < .001$, $\eta_p^2 = .47$, because of increased distributional skewing in the assembled cursive condition.

Error Rates

A comparison of error rates (Table 7) between the computer print and assembled cursive conditions revealed a reliable effect of script, $F_S(1, 51) = 29.01$, $p < .001$, $\eta_p^2 = .36$, with higher error rates in the cursive condition. We also observed a reliable regularity effect, $F_S(1, 51) = 120.43$, $p < .001$, $\eta_p^2 = .70$, and an imageability effect, $F_S(1, 51) = 98.37$, $p < .001$, $\eta_p^2 = .66$. Imageability interacted significantly with script, $F_S(1, 51) = 6.41$, $p < .05$, $\eta_p^2 = .11$, with increased errors to low-imageability, cursive words. Finally, we observed an overall regularity × Imageability interaction, $F_S(1, 51) = 23.64$, $p < .001$, $\eta_p^2 = .32$, as in Strain et al. (1995).

Discussion

Experiments 4A and 4B followed our expectations regarding top-down influences in disambiguating noisy, handwritten words. The imageability effect, while present in the computer print condition, was magnified in the cursive condition. This finding was validated in the ex-Gaussian analyses, showing that mean RT differences between high- and low-imageability words were attributable only to distributional shifting, not skewing. Examination of the effect sizes (η_p^2) for the μ parameter revealed that imageability effects were considerably increased from the print to the cursive condition. Within the framework of Harm and Seidenberg's (2004) model, a handwritten word should naturally delay resonance forming between orthography and phonology, thus increasing reliance on top-down, semantic information to disambiguate items at the letter level, subsequently producing strong benefits for words with more concrete meanings.

As had been reported previously (Monaghan & Ellis, 2002), when AoA estimates are included as a covariate for imageability, most effects of imageability were reduced or eliminated (for brevity, we did not report the analyses without AoA, but their general pattern can be gleaned from the distributional analyses). In the item analyses for Experiment 4A, we included AoA ratings as a covariate, following Monaghan and Ellis (2002; Cortese & Khanna, 2007). In this case, after effects of AoA had been partialled out, a reliable Regularity × Imageability interaction remained (in the print condition, and it was nearly reliable in the cursive condition). This result is at odds with results from Shibahara et al. (2003). In general, we believe the topic of AoA required more research, especially in the present context: Are novel handwritten words appropriate to delineate on scales for AoA, when they are such unusual forms? To what degree do participants merely use imageability and frequency when providing AoA estimates? We hope to focus on these issues more closely in the domain of handwritten word reading.

General Discussion

The uniting characteristic of the present experiments is the exaggeration of lexical effects under conditions of high physical ambiguity, when participants read handwritten words. Both regularity and feed-forward consistency can be considered variables that influence the translation of orthography to phonology, allowing fast, accurate naming. Imageability can also be considered a translational variable, acting from the top-down, as semantics are translated into phonology. Although computer-generated print yields similar effects, they are small, relative to human cursive. This likely reflects the fact that computer print is prototypical and non-polysemous, with well-defined and clearly demarcated letters.

These pristine forms do not impose the same slow-down that heavily polysemous, often novel-looking handwritten forms do. The human perceptual system is equipped to disambiguate handwritten words; it simply has to rely more heavily on top-down processes, relative to more prototypical word forms. Thus, effects of frequency, consistency, and imageability are all enhanced in reading handwritten words.

The present studies add credence to the suggestion by Manso De Zuniga et al. (1991) that handwritten word forms slow lexical access, allowing typically fast processes to exert greater influences on disambiguating the stimuli. Manso De Zuniga et al. observed increased frequency effects in handwritten word recognition; this held true in our replication and extension of Jared (2002; and also of Seidenberg et al., 1984). Regularity effects remained relatively invariant across word forms, but frequency interacted with script, with a larger effect in the recognition of human cursive. In our replication and extension of the Stone et al. (1997) bidirectional consistency experiment, the cursive condition showed a dramatic increase in the (previously minimal) influence of feed-forward consistency, in both lexical decision and naming. Feedback consistency effects showed the same pattern in naming, increasing in the human cursive condition. The same magnification of effects occurred in our replication and extension of Strain et al. (1995): Given handwriting, the benefit of high imageability more than doubled, relative to the computer print condition (although these effects were attenuated when AoA was added as a covariate).

Although our discussions have focused mainly on RTs, it is important to note that the error data were equally compelling. In their examination of semantic effects in word naming, Strain et al. (1995) focused heavily on error rates, consistently observing robust effects of regularity, imageability, and their interaction. The error rates in Experiments 4A and 4B perfectly replicated their results. In similar fashion, the error analyses for our replication and extension of Stone et al. (1997) were equally informative, producing reliable effects of FF and FB consistency and their interaction, which again increased in the cursive condition.

Clearly, more research is needed on handwritten word recognition, which has the potential to clarify the importance (or at least magnitudes) of theoretically meaningful effects. The present research lays the groundwork for more in-depth studies of the factors involved in reading human script. It is still unclear what role top-down processing plays in the recognition of these ambiguous forms. For example, if handwriting is analogous to speech, then top-down processes should play an important role in the segmentation problem, as they do in speech recognition. In order to truly appreciate such processes, it will likely be necessary to move away from single-word naming, into paradigms containing contextual cues that could potentially disambiguate the input. For example, Becker and Killion (1977) examined the effect of semantic context under different levels of stimulus intensity, finding that context effects were magnified when the input was degraded. We would expect these effects to be magnified even more when the input was not only degraded, but ambiguous, as in handwriting.

Further attention should also be paid to the role of implicit memory in handwritten word recognition. Brown and Carr (1993) found that repetition priming effects for printed words did not vary with prime script, but handwritten items benefited greatly from same-script primes, suggesting that surface forms are encoded for atypical scripts. They concluded that, "when the target was in the less familiar and inherently more variable medium of handwriting, an extended history of reading experience with that particular string of letters became important for getting repetition benefit (p. 1,286)." These same effects are observed in the auditory modality, with same-voice repetitions leading to stronger priming benefits (Cole, Coltheart, & Allard, 1974; Goldinger, 1996). In addition to experience with specific letter strings, experience with a handwriting style over time should generalize, in a fashion similar to *font tuning* (Walker, 2008), wherein participants use implicitly encoded script regularities to inform perception of subsequent, new letter strings. It may be the case that handwritten word perception employs a more exemplar-based encoding and recognition strategy, as suggested by Marsolek (2004).

The present study, along with Manso De Zuniga et al. (1991), represents a relatively unique examination of visual word recognition. Handwritten stimuli, although physically noisy, are a more ecologically valid form for research, potentially tapping human faculties that only weakly influence the reading of synthetically generated stimuli. Beyond influences on lexical processing, perception of handwriting may be interesting for other reasons. For example, Hellige and Adamson (2007) found that the right-hemisphere of the brain contributes more to the recognition of handwriting than print. In addition, research by Longcamp, Tanskanen, and Hari (2006) suggests that motor areas of the brain may play a functional role in the recognition of handwritten words, just as the motor theory of speech perception accounts for the recognition of spoken words (Galantucci, Fowler, & Turvey, 2006; Liberman & Mattingly, 1985). And, while handwriting may not provide a window to the personality, as graphologists contend, it does contain visual subtlety that can impart unwritten meaning to the reader. Loewenthal (1975) claimed that participants could manipulate their handwriting to convey predetermined personality traits, but he was likely overstating his case. Although the inferred information was framed as a personality trait, it was more likely some emotional characteristic, such as anger or happiness. Emotional content could be added to text to convey greater meaning than the words alone, just as voice inflection can indicate a speaker's underlying emotional state. Indeed, the perceptual richness of handwritten stimuli could be likened to that of spoken words. If a study of spoken word recognition utilized only computer-generated voices, it would likely be criticized for lacking ecological validity. Perhaps studies of printed word perception could benefit from greater application of natural materials.

References

Allport, D. A., & Funnell, E. (1981). Components of the mental lexicon. *Philosophical Transactions of the Royal Society of London: Series B, Biological Sciences, 295,* 397–410.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 340–357.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133,* 283–316.

Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: Beyond measures of central tendency. *Journal of Experimental Psychology: General, 128,* 32–55.

Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language, 59,* 495–523.

Becker, C. A., & Killion, T. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 389–401.

Besner, D., & Johnston, J. C. (1989). Reading and the mental lexicon: On the uptake of visual information. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 291–316). Cambridge, MA: MIT Press.

Brown, J. S., & Carr, T. H. (1993). Limits on perceptual abstraction in reading: Asymmetric transfer between surface forms differing in typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1277–1296.

Cole, R. A., Coltheart, M., & Allard, F. (1974). Memory of a speaker's voice: Reaction time to same- or different-voiced letters. *Quarterly Journal of Experimental Psychology, 26,* 1–7.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100,* 589–608.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Corcoran, D. W., & Rouse, R. O. (1970). An aspect of perceptual organization involved in reading typed and handwritten words. *Quarterly Journal of Experimental Psychology, 22,* 526–530.

Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words. *The Quarterly Journal of Experimental Psychology, 60,* 1072–1082.

Cortese, M. J., & Khanna, M. M. (2008). Age of acquisition ratings for 3,000 monosyllabic words. *Behavior Research Methods, 40,* 791–794.

Duffy, S. A., & Pisoni, D. B. (1992). Comprehension of synthetic speech produced by rule: A review and theoretical interpretation. *Language & Speech, 35,* 351–389.

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin, 125,* 777–799.

Fowler, C. A. (2005). Parsing coarticulated speech in perception: Effects of coarticulation resistance. *Journal of Phonetics, 33,* 199–213.

Furnham, A., & Gunter, B. (1987). Graphology and personality: Another failure to validate graphological analysis. *Personality and Individual Differences, 8,* 433–435.

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review, 13,* 361–377.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 674–691.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1166–1183.

Goldinger, S. D., Pisoni, D. B., & Luce, P. A. (1996). Speech perception and spoken word recognition: Research and theory. In N. J. Laas (Ed.), *Principles of experimental phonetics* (pp. 277–327). St. Louis: Mosby.

Graham, S., Harris, K. R., & Fink, B. (2000). Is handwriting causally related to learning to write? Treatment of handwriting problems in beginning writers. *Journal of Educational Psychology, 92,* 620–633.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103,* 518–565.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111,* 662–720.

Hellige, J. B., & Adamson, M. M. (2007). Hemispheric differences in processing handwritten cursive. *Brain and Language, 102,* 215–227.

Inhoff, A. W. (1991). Word frequency during copytyping. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 478–487.

Jared, D. (2002). Spelling-sound consistency and regularity effects in word naming. *Journal of Memory and Language, 46,* 723–750.

Jared, D., McRae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language, 29,* 687–715.

King, R. N., & Koehler, D. J. (2000). Illusory correlations in graphological inference. *Journal of Experimental Psychology: Applied, 6,* 336–348.

Lacouture, Y., & Cousineau, D. (2008). How to use MATLAB to fit the ex-Gaussian and other probability functions to a distribution of response times. *Tutorials in Quantitative Methods for Psychology, 4,* 35–45.

Lacruz, I., & Folk, J. R. (2004). Feedforward and feedback consistency effects for high- and low-frequency words in lexical decision and naming. *Quarterly Journal of Experimental Psychology, 57*A, 1261–1284.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74,* 431–461.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21,* 1–36.

Loewenthal, K. (1975). Handwriting and self-presentation. *Journal of Social Psychology, 96,* 267–270.

Longcamp, M., Tanskanen, T., & Hari, R. (2006). The imprint of action: Motor cortex involvement in visual perception of handwritten letters. *NeuroImage, 33,* 681–688.

Lorette, G. (1999). Handwriting recognition or reading? What is the situation at the dawn of the 3rd millennium? *International Journal on Document Analysis and Recognition, 2,* 2–12.

Manso De Zuniga, C., Humphreys, G. W., & Evett, L. J. (1991). Additive and interactive effects of repetition, degradation, and word frequency in the reading of handwriting. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 10–33). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Marsolek, C. J. (2004). Abstractionist versus exemplar-based theories of visual word priming: A subsystems resolution. *The Quarterly Journal of Experimental Psychology, 57*A, 1233–1259.

Massaro, D. W., & Jesse, A. (2005). The magic of reading: Too many influences for quick and easy explanations. In T. Trabasso, J. Sabatini, D. W. Massaro, & R. C. Calfee (Eds.), *From orthography to pedagogy: Essays in honor of Richard L. Venezky* (pp. 37–61). Mahwah, NJ: Lawrence Erlbaum Associates.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

Monaghan, J., & Ellis, A. W. (2002). What exactly interacts with spelling-sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 183–206.

Neter, E., & Ben-Shakhar, G. (1989). The predictive validity of graphological inferences: A meta-analystic approach. *Personality and Individual Differences, 10,* 737–745.

Peereman, R., Content, A., & Bonin, P. (1998). Is perception a two-way street? The case of feedback consistency in visual word recognition. *Journal of Memory & Language, 39,* 151–174.

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review, 114,* 273–315.

Pisoni, D. B., & Luce, P. A. (1987). Acoustic-phonetic representations in word recognition. *Cognition, 25,* 21–52.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

Pressler, M. W. (2006, October 11). Handwriting is on the wall [electronic version]. *Washington Post.* Retrieved from www.washingtonpost.com

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89,* 60–94.

Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General, 110,* 474–494.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh: Psychology Software Tools Inc.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 523–568.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23,* 383–404.

Shibahara, N., Zorzi, M., Hill, M. P., Wydell, T., & Butterworth, B. (2003). Semantic effects in word naming: Evidence from English and Japanese Kanji. *Quarterly Journal of Experimental Psychology, 56*A, 263–286.

Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language, 36,* 337–359.

Strain, E., & Herdman, C. M. (1999). Imageability effects in word naming: An individual differences analysis. *Canadian Journal of Experimental Psychology, 53,* 347–359.

Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1140–1154.

Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms.* Hillsdale, NJ: Erlbaum.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition, 15,* 181–198.

Van Orden, G. C., & Goldinger, S. D. (1994). Interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1269–1291.

Walker, P. (2008). Font tuning: A review and new experimental evidence. *Visual Cognition, 16,* 1022–1058.

Winer, B. J. (1971). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill. AQ: 3

Ziegler, J. C., Montant, M., & Jacobs, A. M. (1997). The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language, 37,* 533–554.

Ziegler, J. C., Petrova, A., & Ferrand, L. (2008). Feedback consistency effects in visual and auditory word recognition: Where do we stand after more than a decade? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 34,* 643–661.

Received May 13, 2008
Revision received December 29, 2009
Accepted January 4, 2010 ■