JOURNAL OF CONSUMER PSYCHOLOGY, *17*(1), 49–58

The SMAART Scale: A Measure of Individuals' Automatic Access to Secondary Meanings in Polysemous Statements

Claudiu V. Dimofte
Georgetown University

Richard F. Yalch
University of Washington

Psycholinguistic and neuropsychological research shows that individuals differ in their ability to access the multiple meanings implied by polysemous expressions. Drawing on these studies, a novel, computer-based measure of automatic access to secondary meaning (SMAART) is developed to distinguish individuals more likely to access only a single, immediately available meaning from those accessing multiple meanings. The new measure is found to be reliable and distinct from several established measures assessing higher level verbal abilities such as the verbal SAT. Several experimental studies demonstrate the scale's usefulness for predicting who is most susceptible to the priming effects of the secondary meanings contained in the polysemous headlines in consumer-oriented communications.

> *"Without the metaphor system, there could be no philosophizing, no theorizing, and little understanding of our everyday personal and social lives. But the operation of this vast system […] is largely unconscious."*
>
> (Lakoff, 1995, p. 229)

In 2005, the governor of the state of New Jersey enlisted the help of Lippincott Mercer, an image consultant firm, to create a new state slogan that would resonate better with residents and tourists. The result of the $260,000 contract was the slogan "*New Jersey: We'll Win You Over*," expected to be applicable in various promotional campaigns (including perhaps those related to the Atlantic City gambling industry). However, it was shelved the day before it was to be officially unveiled by the New Jersey Department of Travel and Tourism because the governor felt that *We'll Win You Over* was a phrase with too many negative connotations (Kidd, 2006)—for one, it reminded him of his dating days and the rejections he failed to win over. Others may have recalled that New Jersey came out the winner when they went there to gamble. New Jersey's slogan problem shows the importance of anticipating and understanding the ways in which various individuals process language when multiple interpretations result from the same linguistic input.

Polysemous statements include metaphors, puns, analogies, and other forms of speech that have multiple interpretations. For example, the expression *Lawyers are sharks* can be understood literally as claiming that lawyers are dangerous creatures that swim in salt water, and figuratively as arguing that lawyers are aggressive individuals who prey on clients or defendants. In such metaphorical contexts, the figurative interpretation is the intended meaning, whereas the literal meaning is something of a syntactic by-product. In general terms however, one meaning is more apparent and immediately available, whereas another is somewhat "hidden" and more difficult to access. Widely used in persuasive communications, one observes polysemy in brand slogans (Michelin's statement that *You Have a lot Riding on Your Tires*), political rhetoric (President Nixon's 1969 launching of a *war on drugs*), and legal arguments (when *the defendant was caught red handed*). Despite their popularity, the communication effectiveness of polysemous statements appears to be less than expected (see Sopory & Dillard, 2002) and not always understood, as the New Jersey slogan example shows.

McQuarrie and Mick (1996) argued that accounting for and explaining individuals' responses to nonliteral

Correspondence should be addressed to Claudiu V. Dimofte, The McDonough School of Business, Georgetown University, Washington, DC 20057, USA. E-mail: dimofte@msb.edu

speech in general requires a more comprehensive view of moderating variables. In these authors' view, these may include individual difference variables such as need for cognition, tolerance for ambiguity, stimulation level, or a specific propensity to respond to figurative language (also see Yarbrough, 1991). Neuropsychological data confirm the existence of individual differences in terms of metaphor comprehension. Burgess and Simpson (1988) proposed that automatic processing of ambiguous meanings occurs in both brain hemispheres, but the left hemisphere has a particularly important role in terms of controlled processing of meanings. Results of PET (positron emission tomography) studies (e.g., Jonides, Smith, Marshuetz, Koeppe, & Reuter-Lorenz, 1998) have found that the left inferior frontal gyrus is activated in verbal working memory tasks, and age-related differences in terms of verbal working memory have also been confirmed by PET research (Jonides et al., 2000). Indeed, Giora, Zaidel, Soroker, Batori, and Kasher (2000) found that patients with brain damage in the left hemisphere showed a significant negative correlation between lesions and scores on particular sarcasm and metaphor comprehension tests. However, lesions in the right hemisphere did not correlate with either test performance. Along similar lines, Ramachandran (2005) recently studied patients with left angular gyrus defects and found that, when asked to explain the deeper meaning of a series of metaphors, these patients always took the phrases literally. However, patients with lesions in different areas of the brain correctly construed the deeper meaning of these metaphors.

Understanding the fundamental sources of individual differences in figurative language comprehension is the goal of the present research. Although neuropsychological or PET techniques have identified physiological sources for figurative language processing differences, they are impractical tools for marketers who are interested in directly assessing these individual differences among consumers. To overcome this problem, we develop in this article a self-administered (computer-based) test that can assess individual differences in metaphoric language processing. This measure of automatic access to meaning (Secondary Meaning Access via the Automatic Route Test—hereinafter SMAART or SMAARTS when referring to the test score) is developed by adapting and building upon a standard sentence verification procedure (see Glucksberg, Gildea, & Bookin, 1982; McCloskey & Glucksberg, 1979). After describing the development process, the measure's usefulness is evaluated by testing its predictive accuracy regarding consumers' implicit processing of polysemous statements. Theoretical implications relative to several unsettled issues regarding how consumers extract meaning from expressions that have multiple interpretations are finally discussed. We begin by presenting the conceptual framework underlying the SMAART scale.

BACKGROUND

Human communications often rely on inferences and implications. Bransford and Franks (1971) argued that what is stored in memories of conversations or messages is the gist of what is said rather than the exact surface form. The extraction of gist seems to be automatic, allowing for the subsequent recall of semantic meaning even as the surface form quickly fades from memory. In the case of a multiple-meaning utterance, the inherent ambiguity of the expression is likely to hinder the process of appropriating a specific meaning from the communication. For example, taking marketplace rumors with a grain of salt is something most readers will figuratively undertake, although the process is unlikely to produce thirst. As both literal and figurative meanings are possible and sometimes equally fitting, consumers must go beyond ordinary language processing to comprehend polysemous phrases. Explaining how this occurs has motivated much cognitive and psycholinguistic research, although there is no agreement yet in terms of the precise processing mechanism behind the comprehension of polysemy.

Early studies of literal and figurative language processing from linguistics and cognitive psychology research postulated the priority of access to literal meanings. These "standard pragmatic models" of discourse comprehension (Giora, 1999) assumed that the initial activation of the literal is mandatory, such that individuals would always access and evaluate the literal meaning of a sentence first (Searle, 1969). However, Verbrugge (1976) claimed that it was erroneous to believe that the literal meaning is the first to be accessed merely because of its supposed cognitive simplicity, and that other factors (most importantly context) also play a role. Several researchers then proposed a "direct access" model of meaning acquisition.

Glucksberg et al. (1982) demonstrated that the comprehension of metaphors is automatic in the sense that individuals cannot ignore the figurative meanings even when directed to only evaluate the literal meanings. Thus, judging the literal truthfulness of the polysemous statement *Some surgeons are butchers* took significantly longer than the same judgment for statements that were literally true (*Some birds are eagles*) or literally false (*Some birds are apples*). Presumably, an automatic generation of the figurative meaning of the metaphoric statements and the required suppressing of their figurative but not literal truthfulness accounts for the slower response time. Glucksberg et al. (1982) further illustrated the automatic access/suppression of figurative meaning process by showing that metaphoric statements not easily recognized as having figurative meanings do not interfere in the processing of the literal meanings. Thus, judging the truthfulness of the statement *All surgeons are butchers* took significantly less time to judge as false than when the statement was *Some surgeons are butchers*. It appears that automatic access to secondary

meanings occurs when statements have secondary meanings and when the comprehender is stimulated to consider these meanings by the wording of the statement.

The present research proposes that individuals may differ in their spontaneous tendency to generate and consider figurative interpretations of polysemous statements in a manner similar to the way different wordings of polysemous statements appear to activate figurative meaning awareness. That is, some individuals may readily become aware of and consider figurative meanings associated with polysemous statements, whereas others may focus on the literal interpretations. This conjecture is supported by the neuropsychological results reviewed in the introduction that isolate separate brain functions associated with figurative language. Accordingly, we propose that automatic access to multiple meanings occurs more often for some individuals than others, and further that this difference is greater when the polysemous statements contain figurative meanings not immediately accessed by individuals. In some instances, for example when individuals are very familiar with both meanings (i.e., equibiased expressions) or when the context strongly implies a given meaning, processing should be similar. However, in situations where this is not the case, this individual difference in polysemous statement processing should be observed. This account underlies a proposed new measure of automatic access to meaning that is described next.

THE SMAART SCALE—A SPEED TEST

Procedure and Scoring

The Secondary Meaning Access via the Automatic Route Test assesses automatic access to secondary meaning by observing whether it interferes with judging the literal veracity of a polysemous statement. It is adapted from a standard test in cognitive psychology (McCloskey & Glucksberg, 1979). The test involves three blocks: one for learning the key assignments (e.g., hitting the "Q" key for true and the "P" key for false sentences), one for practice sentences, and one for test sentences. In the procedure, subjects are requested to verify the literal truth of sentences of the type *Some X are Y*. During the test block, response latencies (accurate to the millisecond) are measured for both random target sentences such as *Some cars are snails* (literally false but figuratively true) and random filler sentences such as *Some flowers are roses* (literally true) and *Some insects are roses* (literally false). The difference between latencies on metaphor and filler sentences (incremental response time) is measured and used as a proxy for automatic comprehension. Errors are assumed to show incapacity to suppress metaphor interference and are treated as a 3,000 msec penalty, roughly equivalent to a three standard deviation latency increase for that item (underscoring the importance of practice trials, note that such errors occurred extremely rarely and never twice for the same subject). Assuming that interference from the available secondary figurative meaning is slowing participants' reaction time, we classify them as having higher automatic access (high-SMAARTS individuals). Those individuals whose average response times for targets and fillers differ very little or not at all are assumed to be less affected by the secondary meanings and are classified as having lower automatic access (low-SMAARTS). Recent research suggests that our measure is likely to adequately capture this ability (see Page, Locke, & Trio, 2005, for an argument along similar lines in the context of ironic processing theory).

There are several variables related to speed that one must keep in mind for a test like SMAART. Although the task speed instructions do not force subjects to perform at any set speed, they do encourage rapid responses that allow automatic reactions to emerge before the engagement of significant conscious processing resources. Although a subject's preferred rate of response can either be specific to the task at hand or reflect a more general personality trait, the speed instructions should not produce changes in individual differences. Moreover, note that—as a latency difference measure—SMAART is insensitive to variability in terms of subjects' base rate of response (i.e., a subject will be similarly slow/fast in hitting the keys for both target and filler sentences and the difference measure will parse out individual speed).

Participant Sample

Essential in scale development is to have a sufficient number of participants: At least 200 normative subjects is a rule of thumb for a stable analysis (Guadagnoli & Velicer, 1988), and the SMAART pilot test employed close to 1,500 participants. The sample used college students enrolled in introductory marketing classes at two major American universities, under the universal assumption that adult cognitive processes involved in language processing are well established by this age. The sample exhibited significant variability in terms of several variables, including verbal ability (i.e., SAT Verbal scores between 300 and 790).

Test Items: Selection and Analysis

The choice of items originated in previous work by Glucksberg et al. (1982) and McElree and Nordlie (1999). The latter reference in particular provided a detailed list of 720 items (i.e., sentences) grouped into three categories: figurative, literal, and nonsense strings—mapping onto our literally false/figuratively true, literally true, and literally false sentences. One hundred items from this source and from the present authors themselves constituted the first iteration of the item pool. Three independent judges evaluated these items in terms of how representative they were for the three

sets, their appropriateness for the subject population, and the need to provide a test of reasonable length (consideration was also given here to the amount of necessary practice blocks). A subset of 43 items was eventually agreed upon: 15 practice block items and 28 test block items (including 15 target and 13 filler sentences—see Appendix for details). Corrected item-total correlations for both target and filler sentences were used as the discrimination index and provided support for the initial item choice.

Methodological heterogeneity is an important aspect of scale development that can mimic content homogeneity by causing correlations between items to cluster. While such outcome may signal issues of concern in the case of scales with items keyed in different directions, it can also hint at the measurement of more than one factor—in this case it should describe the existence of two types of items (polysemous and literal). Indeed, as predicted by psycholinguistic theory, polysemous items were the most difficult to answer (M_{Poly}=1638.04, SD=607.69), followed by literal false (M_{LitF}=1513.23, SD=534.90) and literal true items (M_{LitT}=1488.77, SD=527.21).[1] In subsequent paired t tests all three comparisons emerged significant ($p < .05$). A similar picture emerges when looking at items' difficulty index, as the percentage of examinees who correctly answered all the items also varied by item type (70.67% for figurative and 98.99% for literal items).

Finally, to look at the internal consistency of each group of items as a pseudo-subscale, alpha values of .71 and .79 were obtained for the polysemous and literal sets respectively, both suggesting satisfactory reliability. This subscale analysis is preferable to a factor analysis of item response times because of the relative similarity of items in terms of syntactic structure and explicit primary meaning access, as well as the relatively narrow response latency distribution. Accordingly, factor analysis is likely to produce numerous cross-loadings and no discriminable or meaningful latent components.

Reliability: Split-Half Method

According to Nunnally and Bernstein (1994), "it is *not* correct to measure the reliability of a speed test in terms of internal consistency (coefficient alpha)" (p. 351). Instead, multiple forms of the test may be devised and inter-correlated, or (as a timesaving alternative) the split-half method may be used on a single form of the test. The correlation between two half-tests (i.e., tests employing half of the original items) from the SMAART scale was .78. For a more meaningful estimate of reliability we employed the correction provided by the Spearman-Brown prophesy formula. Then, the corrected correlation between halves reaches .88. As with many psychological tests, learning occurs with repeated performing of the same test. Whereas the reliability discussion above suggested that an alternative version of the test should offer similar results in terms of subjects' measured ability, the SMAART scale does not allow immediately successive results to be stable.

Validity: External Correlates

It is generally required that novel measures show sufficient validity (i.e., the confidence that the domain of content is indeed as conceptually intended), and (despite some idiosyncrasies) speed-tests are no exception. In this spirit, our measure should predictably correlate with alternative measures capturing the same domain of interest/construct (convergent validity), as well as not correlate with those that (although perhaps similar) capture conceptually distinct domains or constructs (divergent validity). As theorized above, the SMAART is proposed to measure an individual's ability to automatically access both meanings of a polysemous expression. The cognitive psychology literature has speculated that automatic access is often facilitated by an inherently larger working memory capacity (see Carpenter, Miyake, & Just, 1994; Daneman & Merikle, 1996; Just & Carpenter, 1992). If this is the case, the convergent validity requirement translates into the need for a significant correlation to be found between subjects' SMAARTS and their performance on a measure of working memory capacity. We developed and used a computer-adapted version of Daneman and Carpenter's (1980) Reading Span Test that requires participants to remember for subsequent recall the last words of a series of 13- to 16-word sentences (for details on the procedure, see Daneman & Carpenter, 1980). Critical for the conceptual account behind the proposed working memory account, this Reading Span Test and SMAARTS is indeed correlated but weakly (r=.12, N=328, $p < .04$).

The divergent validity requirement addressed the automatic aspect of the ability being measured. Here, two alternative measures were looked at: first, the SAT Verbal score. Daneman and Hannon (2001) demonstrated a correlation between working memory and SAT performance, suggesting that some common processes such as automatic access may affect one's score. However, although the SAT's time constraint encourages quick responses, the time taken for each question is not recorded and test takers benefit from higher level processes. Furthermore, it is unlikely that there are many polysemous statements in the SAT, as it is a multiple-choice test with only one correct answer per question. Consistent with these processing differences, the correlation between SMARTS and reported SAT verbal scores approached zero (r=–.003, N=184, ns). Second, the Polychronic Attitude Index (PAI; Kaufman, Lane, & Lindquist,

[1]Note that the average incremental response time for our sample is 137.87 msec (M_{Poly}— mean(M_{LitF}, M_{LitT})). In other words, the zero of our (relative) scale is in fact 137.87. Confirming the validity of SMAART and the fact that this ability is normally distributed, the median split was around 140 msec across all our studies (some unreported here).

1991) measured individuals' self-reported belief in their ability to engage in multiple concomitant behaviors (i.e., to consciously multitask). As the automatic access to meaning is beyond conscious control, this measure should only correlate slightly with SMAARTS, a fact indeed observed (r=−.13, N=184, $p < .07$). Overall, these analyses confirm that the SMAART scale assesses an automatic verbal ability distinct from more conscious verbal abilities and one not likely to be self-reported.

Two issues remain unclear about SMAART. First, the scale admittedly uses a forced choice procedure that is quite dissimilar to processes involved in regular comprehension of polysemous discourse. Second, the primary/secondary meaning distinction is created artificially by task-imposed directions and does not necessarily mimic the same linguistic distinction as it naturally occurs. Typically, meaning acquisition is determined by both the relative base frequency of each meaning and what is suggested by the context in which the statement appears. An example of the latter is that *A grain of salt* is likely to be interpreted figuratively when the context concerns the mass media but literally when the context is food-related. Predictive validity studies reported later will address both of these concerns.

Standardizing Norms

Several guidelines are offered next based on successive scale calibration efforts undertaken with over 1,000 individuals describing an extremely diverse group in terms of race, age, gender, and socio-economic status. Because the original SMAARTS distribution is approximately normal, the effects of scores normalization would be slight, and thus the present standardization norms make use of the raw sample population scores. This procedure was adopted notwithstanding several important caveats that previous researchers (notably McQuarrie & Phillips, 2005) have made related to the perils of analyzing untransformed latency data. In the present case, the latency distributions do not depart from normality (neither for individual items nor for the incremental response time), so log transformations are unnecessary. Furthermore, the lack of a skew can be explained by the nature of the speed-test, wherein the quick responses produce uniformly low mean response times (around 1500 msec) and low variance. In terms of the false positive responses, penalizing errors via increased latencies is a practice adopted is other latency tests as well (e.g., the Implicit Association Test; Greenwald, McGhee, & Schwartz, 1998). Finally, to counteract the emergence of a practice effect within test items, participants are first given several practice trial items that allow for minimal marginal gains in terms of speed of response to the subsequent set of test items. The randomized sequence of test items further mitigates any potential learning problems.

For the purposes of sample splits along subjects' SMAARTS, we recommend the use of a 5% trimmed median of 140 milliseconds. Using weighted average SMAARTS, we observed the following rough percentiles: 5th at –320 msec, 10th at –190 msec, 25th at –25 msec, 50th at 120 msec, 75th at 275ms, 90th at 475 msec, and 95th at 700 msec. Some tests provide percentile-ranking information for each of their different components, but two factors go against such procedure in our case. First, employing this rationale for the polysemous and literal sets of items would provide no interesting conceptual information (except the fact that a lower mean score occurs in the latter group, a fact already mentioned). Second, and more importantly, the separate between-subject comparisons would be biased by individual differences in test-taking strategies or speed of response, problems avoided by the use of the final difference measure.

Finally, issues related to range restrictions should be salient at this point, as the mostly college student sample population employed in the present scale development process was admittedly not perfectly representative (in the statistical sense) of the general population. One could therefore argue that a less educated population would experience less interference from the secondary meaning (i.e., less automatic access to it) and this would drive down the average statistics for SMAARTS. However, our physiological-based theoretical explanation does not easily warrant such inference. Moreover, across numerous samples and testing instances the bell-shaped curve emerged robustly for the response time distribution, suggesting that the ability that SMAART measures is indeed distributed normally in the population.

Other Issues

In many speed tests issues of guessing are pertinent, especially as many low-ability subjects tend to engage in such behavior. Two mechanisms were employed in order to keep the SMAART scale devoid of such problems. First, the test instructions offered participants clear guidelines in terms of responding to items, while the practice block provided feedback in terms of the correct response and explained why the alternative was incorrect. Second (as mentioned above), the possibility that participants engaged in guessing during the test phase would surely result in high error rates (recorded as such by the software). In these cases (i.e., subjects with more than 30% of answers wrong), the involvement or motivation (see Chan, Schmitt, DeShon, Clause, & Delbridge, 1997) was deemed sufficiently low as to bias the result in a manner that would invalidate their score (for example, repeatedly failing to acknowledge that items such as *Some birds are eagles* are literally true sentences clearly suggests not paying attention or random guessing). Consequently, these subjects were dropped from the analyses (note that this occurred in less than 3% of the cases). The low number of errors of commission observed suggests that sophisticated algorithms to correct for guessing need not be employed here.

Predictive Validity: Study 1a

To this point, it has been established that individuals' ability to judge the literal veracity of individual polysemous statements appears affected by their automatic access to a competing figurative meaning. Beyond observing this phenomenon, it is important to demonstrate that this individual difference has consequences for their processing of communications. To demonstrate this, individuals were presented with written communications containing a polysemous statement with a secondary meaning implying something about their experience. It was expected that individuals' response to the statement would vary depending on their position on the SMAART scale, such that judgments made by high access individuals would be affected by the secondary meaning more than judgments made by the low access individuals. This study is essentially a priming experiment along the lines of Bargh, Chen, & Burrows (1996), who showed that the unconscious activation of specific social stereotypes has direct priming effects in terms of subsequent stereotype-confirming behavior. In a similar vein, the automatic access to secondary meaning by high-SMAARTS participants should prime specific reactions to the presented stimuli (different from those of low-SMAARTS participants).

Stimuli. A brief (fewer than 300 words) essay was developed addressing the topic of winter sports in the mountain resort of Aspen. Across four PowerPoint® slides, summary background information was given on the resort facilities, tourists, and predominant types of skiing on the local slopes. Depending on condition, the title of the article (across all slides) and its last line were either *Going Fast in Aspen* (literal) or *Going Downhill Fast in Aspen* (polysemous). Whereas the primary meaning of the latter tag line addresses the fun, speedy runs possible on Aspen's slopes, a secondary meaning suggests the worsening conditions in the town resort of Aspen. As this meaning is in direct contrast with the actual content and tone of the article, the study allows for a very conservative test of the automatic meaning access account.

Subjects and measures. One hundred and five college students took part in the study in return for partial credit toward fulfilling the requirements of an introductory marketing course. After being exposed to the essay, explicit measures were collected. Participants were required to perform an Implicit Association Test (IAT; Greenwald et al., 1998) looking for implicit associations between two similar winter resorts (Aspen and Vail) and the attributes representing the categories of "improve" and "worsen" (e.g., advance and enhance versus decline and deteriorate, respectively). Finally, participants' SMAARTS were collected. The post hoc assessment of SMAARTS by meaning condition assignment showed a very even distribution (high-SMAARTS to low-SMAARTS proportions of 27/27 in the literal and 26/25 in the polysemous condition).

Procedure. Participants were exposed, on computer screens, to the article mentioned above and were requested to read it at their own pace, under the pretext of a subsequent memory test. Returning to the article was impossible upon finishing the reading task. Subsequently, the dependent measures were collected (explicit and implicit) and participants performed SMAART. No individual reported any suspicion as to the true nature of the experiment. Finally, they were debriefed and thanked.

Results. Planned contrasts between high-SMAARTS participants' response to the explicit item assessing the quality of skiing in Aspen showed that they perceived it to be significantly inferior in the polysemous condition (M_{Lit}=6.37, M_{Poly}=5.80), $t(50)=2.18$, $p < .05$. Moreover, a similar result emerged in terms of the explicit comparison of Aspen and Vail for these high-SMAARTS individuals, on a scale of 1 (*strongly prefer Vail*) to 7 (*strongly prefer Aspen*), M_{Lit}=5.15, M_{Poly}=4.80, $t(50)=1.83$, $p=.07$ (see Table 1 for all means). Finally and importantly for the automatic access account, the same planned comparison using the IAT scores showed significantly stronger implicit associations of Aspen and "declining" (using the D-measure of IAT response latencies—see Greenwald, Nosek, & Banaji, 2003—where higher values suggest a closer Aspen-declining association, M_{Lit}=–.22, M_{Poly}=–.02), $t(50)=-3.13$, $p < .01$. No differences were observed between conditions for low-SMAARTS participants (M_{Lit}=–.23, M_{Poly}=–.14), $t(50)=-1.56$, *ns*. Participants' familiarity with skiing or the resort of Aspen did not impact the results.

Predictive Validity: Study 1b

Study 1a demonstrates that when a polysemous statement has a figurative secondary meaning, judgments by individuals with high SMAARTS are more influenced by the secondary

TABLE 1
Means and Standard Errors for Studies 1a, 1b, and 2

	Low-SMAARTS		*High-SMAARTS*	
	M	*SE*	*M*	*SE*
Study 1a: Quality of skiing				
Literal: *Going fast in Aspen*	5.63	0.26	6.37	0.13
Polysemous: *Going downhill fast in Aspen*	5.81	0.20	5.80	0.23
Study 1b: Paragraph spacing				
Polysemous: *Reading between the lines*	1.46	0.06	1.68	0.07
Study 2: Attitude toward ad				
Literal: *Unlike any other*	5.46	0.20	5.77	0.20
Polysemous: *No one comes close*	5.38	0.26	5.22	0.19
Literal: *The obvious choice*	4.68	0.23	4.52	0.23
Polysemous: *The natural choice*	4.65	0.24	4.90	0.19

meaning compared to those individuals with low SMAARTS. However, the secondary meaning of a polysemous statement may not always be a figurative one. Consequently, a second study was conducted to see whether the two groups also differ when the secondary meaning is literal instead of figurative. Thus, in Study 1b, subjects were presented with a polysemous statement containing a figurative primary meaning and a literal secondary meaning based on relative frequency in common language. Susceptibility to the primed judgment suggested by the secondary meaning was taken as evidence that subjects were influenced by it.

Stimuli. A brief (fewer than 500 words) essay was developed addressing the topic of regulating the labels of bottled water brands in the United States. Summary background information was given on sources of such water, percentages of each source out of the total, specific standards of cleanliness, and some problems the FDA had encountered in trying to enforce these standards. The title of the article and its last line were "Reading between the lines of bottled water labels." Whereas the primary meaning of this tag line addressed the need to go beyond appearances when it comes to evaluating the quality of the product, a secondary meaning suggests the actual search for letters in between the rows of a piece of writing. The likelihood of finding such letters should depend on whether the paragraph provides adequate line spacing. In the respective article, line spacing was set at slightly below 1.5 lines (exactly 15 points) to create an ambiguous and thus suggestible distance.

Subjects and measures. One hundred sixty-one college students (all native English speakers) took part in the study in return for partial credit toward fulfilling the requirements of an introductory marketing course. Hidden among a series of irrelevant questions about the topic, the dependent measure included participants' recollection of the article's paragraph spacing on a scale of 1 (*single-space*) through 4 (*larger than double*). They then performed the previously developed test, whereby response latencies classified them as high or low-SMAARTS individuals.

Procedure. Participants were provided with a sheet of paper with the printed article and requested to read it at their own pace, under the pretext of a subsequent memory test. Papers were collected upon finishing the reading task. After 15 minutes of filler tasks, the dependent measure and participants' SMAARTS were collected. No individual reported any suspicion as to the true nature of the experiment. Finally, they were debriefed and thanked.

Results. The post hoc assessment of paragraph spacing showed that high-SMAARTS participants engaged in significant overestimation ($M_{\text{lo-SMAARTS}}$=1.5, $M_{\text{hi-SMAARTS}}$=1.7), $t(159)$=–2.37, $p < .02$ (see Table 1 for all means). With the treatment identical across the two groups, it is apparent that the secondary (here, literal) meaning of the tag line was accessed by the high-SMAARTS subjects, and our measure managed to tap into processes occurring via an automatic route. These results suggest that SMAARTS is an adequate predictor of individuals' [in]ability to automatically access the secondary meaning of a polysemous communication.

To address the possibility that high-SMAARTS individuals tend to make higher spacing estimates than low-SMAARTS individuals no matter the prompt, a small replication study (45 subjects) was conducted. A control condition with the headline "Reading the lines…" was compared to the polysemous "Reading between the lines…" using the similar dependent measure of estimating the spacing of the text. Consistent with the prior results, 67% of the high-SMAARTS individuals exposed to the polysemous headline made high estimates versus only 26% across the other three conditions combined (χ^2=4.1, $p < .05$, df=1, N=45). Further, participants' open-ended thoughts yielded no differences across all four groups, supporting the view that the processing mechanism involved is of an automatic nature and does not involve explicit access to meaning.

Predictive Validity: Study 2

Stimuli. Four promotional messages were developed: 2 for Minute Maid juice and 2 for Mercedes-Benz automobiles. For Minute Maid, one message employed the single-meaning slogan *The Obvious Choice*, while the other used the polysemous version *The Natural Choice*. For Mercedes-Benz, the first message employed the single-meaning slogan *Unlike Any Other*, while the second used the polysemous version *No One Comes Close*. Note that both polysemous slogans have secondary meanings that are literal (as in preservative-free for Minute Maid—a positively valenced meaning—and as in price-driven inapproachability for Mercedes-Benz—a negatively valenced meaning).

For both brands, the stimuli included several pictures of the product, the slogan, and the brand logo. It was expected that upon exposure to both messages, high-SMAARTS individuals would exhibit higher attitudinal scores in the polysemous case for the Minute Maid message (due to their automatic access to the secondary, positively valenced meaning) and lower scores in the polysemous case for the Mercedes-Benz message (due to automatic access to the secondary, negatively valenced meaning).

Participants and measures. Thirty-two college student participants (all native English speakers) took part in the study in return for the chance to win one of four $10 prizes. A multiple-item measure for message attitudes was employed as dependent variable. It consisted of 4 seven-point semantic differential items and had good reliability (Cronbach's $\alpha >$ .85). An open-ended question asking participants to describe in their own words the idea conveyed by the message was coded for mentioning the nonpreservative/expensiveness

aspect or not. The final design was a mixed one, with within-subject exposure to brand and slogan meaning and between-subjects measurement of SMAARTS.

Procedure. Participants were first exposed to two messages, one for each brand, and 10 days later to the complementing ads (exposure order for brands and slogan meanings was randomized across subjects). Thus, a typical subject might see a polysemous Mercedes ad and a single meaning ad for Minute Maid on day one. Ten days later, they would see a single meaning ad for Mercedes and a polysemous ad for Minute Maid. Dependent measures were collected after each message exposure. After the two exposures participants performed the SMAART. At the end of the second experimental session, they were debriefed and thanked.

Results. Pairwise planned contrasts between high-SMAARTS participants' attitudes toward the Minute Maid messages found more favorable reactions toward the ad for those in the polysemous condition than in the single meaning condition ($M_{HighSMAARTS1}$=4.90, $M_{HighSMAARTS2}$=4.52), $t(14)=2.88$, $p < .02$. Results were similar for the Mercedes-Benz message: less favorable attitudes toward the ad for high-SMAARTS individuals in the polysemous condition than in the single meaning condition ($M_{HighSMAARTS1}$=5.77, $M_{HighSMAARTS2}$=5.22), $t(14)=5.44$, $p < .001$ (see Table 1 for all means).

More importantly, these results emerged despite the fact that a binary logistic regression on participants' open thoughts (coded for mentioning the non-preservative/expensiveness aspect or not) revealed that SMAARTS did not moderate explicit access to secondary meaning in the polysemous condition for either brand ($B=.37$, $p=.65$ for Minute Maid and $B=.53$, $p=.54$ for Mercedes-Benz), suggesting that it is automatic and not explicit access to meaning that drives the observed differences.

Discussion. Results suggest that exposure to polysemous slogan messages with positive/negative secondary meanings produces a cumulative/subtractive effect on attitudes among high-SMAARTS individuals. These results cannot be explained by looking at participants' self-reported (thus explicit) meaning access thoughts, a fact suggesting the operation of an automatic processing mechanism among these individuals (as predicted by their SMAART performance).

GENERAL DISCUSSION AND CONCLUSIONS

This research introduced and validated an online speed test (SMAART) that is based on an extant literal truth decision task from cognitive psychology. The measure captures individuals' ability to automatically access the secondary meaning of a polysemous sentence by measuring their success at consciously trying to suppress it. The test is thus similar to lexical decision tasks employing ambiguous words, but an ambiguity *disadvantage* is expected when following instructions (see Rodd, Gaskell, & Marslen-Wilson, 2002 for a related discussion). The scale's uniqueness in assessing automatic verbal abilities compared to higher level ones was demonstrated using comparisons with several standard verbal tests.

Whereas the SMAART was proposed to capture individuals' ability to automatically access secondary meanings of polysemous statements, the test admittedly evaluated this process only for the case of figurative secondary meanings. This ability, we argued, is revealed by a participant's difficulty (or outright inability) to suppress these figurative meanings when the task highlights the literal meanings (via instructions requiring literal truth judgments). Study 1b complemented this one-sided approach by looking at the scale's predictive validity for a polysemous statement with a literal secondary meaning. Results support the contention that, despite the exclusive use of figurative secondary meanings within the test items, SMAART does in fact measure automatic access to secondary meanings in general, be they figurative or literal. It is important to note at this point that this finding is also supported by cognitive psychology research on working memory individual differences. Just and Carpenter's (1992) theory of the way working memory capacity constrains comprehension proposed that knowledge processing and storage are mediated by activation, with the total available amount in working memory varying across individuals. The larger capacity of some individuals allows them to cope better in cases of ambiguity (such as those involved in polysemous expressions), as it apparently permits them to access and maintain multiple interpretations (i.e., literal and figurative). In this sense, high-SMAARTS individuals are similar to those enjoying the benefit of high working memory capacity.

Three studies addressed the predictive validity of the new measure and found it, in two different contexts (verbal only in study 1b, and verbal/pictorial in studies 1a and 2), to successfully predict individual response to communications that make use of polysemous discourse. The fact that the secondary meaning of a polysemous expression is accessed automatically by only some individuals was revealed in study 2 by the occurrence of a cumulative effect of a positive secondary meaning and a subtractive effect of a negative secondary meaning on high-SMAARTS participants' attitudes toward particular polysemous communications. Conversely, low-SMAARTS individuals did not show access to the secondary meaning and therefore not shifted attitudes significantly. Future research is needed to see if more specific instructions (cf. Peters & Nunez, 1999) or more relaxed time limits can benefit this group in terms of accessing more than the immediately available meanings.

It should also be noted that the ability captured by SMAART appears to be distinct from any valence-laden preference or intrinsic individual affinity for figurative language (see Yarbrough, 1991). Our measure assesses an automatic processing mechanism that is related to working memory capacity differences, and is relatively silent on any

inferences this may have for the affective preference for nonliteral language. Along these lines, work by Gerrig and Healy (1983) suggests the fact that meaning access and polysemy appreciation are distinct and unrelated phenomena. Future research is therefore needed to more closely evaluate the cognitive-affective distinction in terms of figurative language processing and its impact on consumer attitudes.

Finally, the SMAART scale makes apparent that numerous cognitive processes that occur beyond conscious awareness have substantive importance for the inferences and choices that individuals make (see Bargh, 2002). One of the studies presented here used the Implicit Association Test to show that individuals exposed to polysemous taglines build automatic associations between these statements' secondary meaning and promoted brands (many times despite not showing paralleling changes in explicit attitudes). It is easy to see how similar effects could occur and would be of significant practical importance in non-marketing situations such as political discourse, the internal communication of organizational behavior norms, the design of social policy campaigns, etc. Future work is therefore needed to more precisely assess the relevance and predictive power of SMAART in such non-consumer contexts where individual differences in the processing of discourse also have immediate implications on attitude formation and decision-making.

ACKNOWLEDGMENTS

This research was supported in part by fellowships to the first author from the University of Washington and Georgetown University. We are grateful for the help and comments provided by Anthony G. Greenwald.

REFERENCES

Bargh, J. A. (2002). Losing consciousness: Automatic influences on consumer judgment, behavior, and motivation. *Journal of Consumer Research, 29*(2), 280–286.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype priming on action. *Journal of Personality and Social Psychology, 71*(2), 230–244.

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 2*(4), 331–350.

Burgess, C., & Simpson, G. B. (1988). Cerebral hemispheric mechanisms in the retrieval of ambiguous word meanings. *Brain and Language, 33,* 86–103.

Carpenter, P. A., Miyake, A., & Just, M. A. (1994). Working memory constraints in comprehension. Evidence from individual differences, aphasia, and aging (pp. 1075–1122). In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics*. New York: Academic Press.

Chan, D., Schmitt, N., DeShon, R. P., Clause, C. S., & Delbridge, K. (1997). Reactions to cognitive ability tests: The relationships between race, test performance, face validity perceptions, and test-taking motivation. *Journal of Applied Psychology, 82*(2), 300–310.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

Daneman, M., & Hannon, B. (2001). Using working memory theory to investigate the construct validity of multiple choice reading comprehension tests such as the SAT. *Journal of Experimental Psychology: General, 130*(2), 208–223.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*(4), 422–433.

Gerrig, R. J., & Healy, A. F. (1983). Dual processes in metaphor understanding: Comprehension and appreciation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*(4), 667–675.

Giora, R. (1999). On the priority of salient meanings: Studies of literal and figurative language. *Journal of Pragmatics, 31,* 919–929.

Giora, R., Zaidel, E., Soroker, N., Batori, G., & Kasher, A. (2000). Differential effects of right and left hemispheric damage on understanding sarcasm and metaphor. *Metaphor and Symbol, 15*(1/2), 63–83.

Glucksberg, S., Gildea, P., & Bookin, H. B. (1982). On understanding nonliteral speech: Can people ignore metaphors? *Journal of Verbal Learning and Verbal Behavior, 21,* 85–98.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464–1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scaring algorithm. *Journal of Personality and Social Psychology, 85,* 197–216.

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin, 103,* 265–275.

Jonides, J., Marshuetz, C., Smith, E. E., Reuter-Lorenz, P. A., Koeppe, R. A., & Hartley, A. (2000). Age differences in behavior and PET activation reveal differences in interference resolution in verbal working memory. *Journal of Cognitive Neuroscience, 12*(1), 188–196.

Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., & Reuter-Lorenz, P. A. (1998, July). Inhibition in verbal working memory revealed by brain activation. *Proceedings of the National Academy of Science, 95,* 8410–8413.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122–149.

Kaufman, C. F., Lane, P. M., & Lindquist, J. D. (1991). Exploring more than 24 hours a day: A preliminary investigation of polychronic time use. *Journal of Consumer Research, 18,* 392–401.

Kidd, L. O. (2006, January 13.). New Jersey's new slogan: "Come See for Yourself." *The Daily Record,* Local News section, A1.

Lakoff, G. (1995). *The Neurocognitive Self in the Science of the Mind.* Oxford, England: Oxford Press.

McCloskey, M., & Glucksberg, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology, 11,* 1–37.

McElree, B., & Nordlie, J. (1999). Literal and figurative interpretations are computed in equal time. *Psychonomic Bulletin & Review, 6,* 486–494.

McQuarrie, E. F., & Mick, D. G. (1996). Figures of rhetoric in advertising language. *Journal of Consumer Research, 22*(4), 424–438.

McQuarrie, E. F., & Phillips, B. J. (2005). Indirect persuasion in advertising. *Journal of Advertising, 34*(2), 7–20.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory (3rd Ed.).* New York, NY: McGraw-Hill.

Page, A. C., Locke, V., & Trio, M. (2005). An online measure of thought suppression. *Journal of Personality and Social Psychology, 88*(3), 421–431.

Peters, W. W., & Nunez, N. (1999). Complex language and comprehension monitoring: Teaching child witnesses to recognize linguistic confusion. *Journal of Applied Psychology, 84*(5), 661–669.

Ramachandran, V. S. (May, 2005). *Role of temporal-parietal-occipital junction and mirror neurons in action-metaphors.* Paper presented at the 17th Annual Convention of the American Psychological Society, Los Angeles.

Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: semantic competition in lexical access. *Journal of Memory and Language, 46*(2), 245–266.

Searle, J. R. (1969). Metaphor. In A. Ortony (Ed.), *Metaphor and Thought*. Cambridge, MA: Cambridge University Press.

Sopory, P., & Dillard, J. P. (2002). The effects of metaphor on persuasion: A meta-analysis. *Human Communications Research, 28*(3), 382–419.

Verbrugge, R. R. (1976). Resemblances in language and perception. In R. Shaw & J. D. Bransford (Eds.), *Perceiving, acting, and comprehending: Toward an ecological psychology.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Yarbrough, D. B. (1991). The reliability and validity of a measure of reported affinity for figurative language. *Creativity Research Journal, 4*(4), 317–335.

APPENDIX

Instructions for Key Assignment Practice Block

We begin with a few trials that will help you get used to the key assignments in today's study.

You will now see the words TRUE or FALSE, in random order.

- push the Q key for true
- push the P key for false

Please hit the appropriate key. (Hit the spacebar to begin)

Instructions for Practice Sentences Block

Very good. You will now see several sentences of the form *Some X are Y*.

You are asked to evaluate these sentences in terms of their literal truth.

For example:

- "Some animals are pigs" is a literally TRUE sentence (indeed, some animals are pigs)
- "Some animals are tulips" is a literally FALSE sentence (indeed, no animal is a tulip)
- "Some people are pigs" is a literally FALSE sentence (though some people are quite sloppy or rude, no human is literally a pig)

Along these guidelines, please evaluate the literal truth of the next sentences as follows:

- push the Q key if the sentence is literally true
- push the P key if the sentence is literally false

The next sentences are for practice purposes.

There will be about two seconds between your response and a new sentence.

Please focus your attention and make your responses as quickly and accurately as you can.

(Hit the spacebar to begin.)

Some flowers are roses.	Some marriages are iceboxes.	Some vehicles are snails.
Some roosters are clocks.	Some toys are dolls.	Some students are sophomores.
Some flowers are eagles.	Some children are angels.	Some rivers are forests.
Some smokers are chimneys.	Some workers are miners.	Some minutes are years.
Some roads are snakes.	Some candles are apples.	Some politicians are democrats.

(ERROR message: ERROR. While some roads are quite twisted, no road is literally a snake. The sentence is literally false.)

Instructions for Test Sentences Block

Very good. The following sentences are for testing purposes.

Once again:

There will be about two seconds between your response and a new sentence.

Please evaluate the literal truth of these sentences as follows:

- push the Q key if the sentence is literally true
- push the P key if the sentence is literally false

Please focus your attention and make your responses as quickly and accurately as you can.

(Hit the spacebar to begin.)

Randomized test items follow.

Literally False and Figuratively True Items	*Literally True Items*	*Literally False Items*
Some surgeons are butchers.	Some birds are eagles.	Some flowers are bees.
Some lawyers are sharks.	Some fish are trout.	Some insects are roses.
Some jobs are jails.	Some trees are oaks.	Some houses are desks.
Some roads are snakes.	Some children are girls.	Some lakes are clothes.
Some pillows are bricks.	Some paintings are portraits.	Some clouds are newspapers.
Some celebrities are gods.	Some liquids are drinks.	Some books are pencils.
Some jokes are bombs.	Some cars are sedans.	Some mountains are lamps.
Some offices are cages.	Some appliances are ovens.	
Some athletes are machines		
Some ideas are gold.		
Some men are mice.		
Some hearts are stone.		
Some words are knives.		