Automatization and Retention of Literacy Skills in Adult Learners

JAMES M. BEBKO, THOMAS RHEE, CARLY McMORRIS, and MAGALI SEGERS

Abstract

Findings from recent efficacy studies comparing literacy program types suggest that struggling adult readers often make limited to moderate gains across varied types of literacy interventions, with no specific approach consistently surpassing others to date. An alternative to comparing program types is to investigate whether there are specific characteristics or skills that vary by individual that can predict higher gains and skill retention across program type. Using an experimental, prospective, longitudinal design, the present study examined the role of automatization (over-learning) of component skills involved in reading during participation in general literacy programs. On average, participants in the study gained the equivalent of one full reading grade-level after participation in programs for six months. The degree of automatization of reading skills was found to be the strongest predictor of gains made during programs; a measure of automatization was also the strongest predictor of subsequent retention of skills, months later at follow-up testing. Implications for adult literacy practitioners and directions for future research related to skill retention are discussed.

Automatization and Retention of Literacy Skills in Adult Learners

It is well acknowledged that poor literacy skills in adulthood are correlated with a number of negative life outcomes. Poor literacy skills can negatively impact employment opportunities, health outcomes, and civic engagement (Vanderberg, Pierce, & Disney 2011). Additionally, adults who are illiterate often experience low self-worth and symptoms of depression, an outcome that has also been documented in older adults (Roman, 2004, Weiss, Fransic, Senf, Heist, & Hargraves 2006). In the United States, limited reading abilities have been associated with poverty, incarceration, and low voter turnout, highlighting adult literacy as a key social issue (National Center for Education Statistics 1993, Kozol 1985). According to the 2003 National Assessment of Adult Literacy (NAAL), 30 million U.S adults are considered to be at the Below Basic level in literacy skills, skills that allow one to search and comprehend printed text (Kutner, Greenberg, Jin, & Paulsen 2006). In Canada there is a similar picture; the 2003 International

Adult Literacy and Skills Survey found that poor literacy skills were associated with health issues, decreased community engagement, and limited employment opportunities (Statistics Canada 2003). Furthermore, 48% of the adult population performs below Level 3 on literacy scales, the level considered reflective of competence in the current knowledge driven economy (Statistics Canada 2003). Generally similar results, with individual variability, have been found in other countries (e.g. New Zealand, Australia, Norway, Switzerland) taking part in the international survey (e.g., Satherley, Lawes, & Sok 2008, Statistics Canada 2003, Walker, Udy, & Pole 1996). Simply put, adult literacy is crucial for full engagement in society, impacting, among others, in social, political and economic areas.

Despite its widespread social impact, adult literacy continues to be a neglected topic in literacy research. In a recent survey of leaders in literacy research, all respondents voted adult literacy as an "extremely cold" research topic area, yet 75% felt that it deserved to be a "hot" topic (Cassidy, Valadez, Garrett, & Barrera 2010). In terms of understanding the instructional and learning processes unique to adult literacy, the field is still in its infancy (Greenberg, Wise, Morris, Fredrick, Rodrigo, Nanda & Pae 2011, Sabatini, Shore, Holtzman & Scarborough 2011). While there is evidence that adults are able to improve their literacy skills (Gombert 1994, Morais, Bertelson, Cary, & Kolinsky 1988), it is unclear to what extent these skills are maintained over time.

There has been interest in determining the types of instructional approaches that would best benefit adult learners and whether these approaches are similar to those used with children. However, it is difficult to evaluate literacy programs due to the diversity of learners (e.g., gender, first language, ethnicity, socioeconomic status, reason for entering the literacy program, etc.), as well as the diversity of the instructional settings and philosophies, including, for example, self-instructional and workplace programs, community colleges, adult high schools or prisons (Sticht & Armstrong 1997, Torgerson, Porthouse & Brooks 2005). There have been several attempts to evaluate efficacy, but because of the inherent diversity of learners as well as inconsistencies in methodologies, the results are equivocal and difficult to interpret and/or generalize to any given literacy program (Torgerson et al 2005). Even within studies, there is considerable variation in outcome, depending on which variables are being measured (Venezky & Sabatini 2002).

Four key literacy concepts are featured in various forms in the majority of intervention programs:

Comprehension -- understanding both the literal and inferential meaning of text

Decoding -- understanding the grapheme-phoneme correspondence of letters and sounds

 Fluency -- the ability to read quickly and accurately

A*utomaticity* -- a sub-process within fluency in which words are recognized accurately and effortlessly (National Research Council, 1998).

Three recent studies have compared the efficacy of approaches using different subsets of these concepts.

Greenberg and colleagues (2011) designed a randomized control study in which 198 struggling adult readers were assigned to one of 5 types of literacy intervention programs, designed to investigate the effectiveness of two main approaches: explicit teaching (e.g., directly teaching how to decode and comprehend) vs implicit teaching (encouraging selection and reading of text that is of individual interest without explicitly teaching component skills). While participants did make gains in their reading abilities, these gains were small and did not differ significantly across the intervention groups. That is, participants in each group made significant gains in their reading abilities as measured by the reading subtests from the Woodcock-Johnson Psychoeducational Battery III (Woodcock & Johnson, 1990), although the size of the gains are considered small at best (.03-.18), using Cohen's (1988) standard conventions that effect sizes between .010 and .30 are small, .30 to .50 medium, and .50 and above large. When a 6-month follow-up was conducted, neither intervention type nor any other variable (age, attendance, entry skills) predicted change in scores over time, and results were mixed with small gains in some areas and small losses in others (D Greenberg, personal communication, November 18, 2011).

Alamprese, MacArthur, Price and Knight (2011) conducted a randomized control field trial in which the treatment groups received programming that focused on spelling and decoding while comparison groups received existing, non-specialized reading instruction (i.e., varied reading instruction that emphasized spelling, vocabulary, and comprehension rather than encoding). Participants in the treatment group made slightly larger gains than the comparison group on one measure of decoding skills, a word attack subtest (effect size = .19); however, there were no differences in gains in word recognition, spelling, fluency, or comprehension. This study also assessed whether participants' background characteristics (place of birth, education, and attendance) were associated with improvement in literary skills. Adult learners who were born outside of the United States (non-native learners) made greater improvements on the majority of reading measures compared to learners born in the country (effect size ranging from .01-.58), regardless of intervention type.

Finally, Sabatini and colleagues (2011) also conducted an intervention efficacy study to determine whether specific kinds of instruction were more

successful in improving the reading skills of adults. Participants were exposed to adapted forms of three intervention programs that have been established for use with children: 1) Corrective Reading, a program that focuses on strengthening reader's grapheme-phoneme correspondences and word recognition; 2) Retrieval, Automaticity, Vocabulary Elaboration – Orthography (RAVE-O), which supplements phonics instruction with a strong emphasis on fluency training; and 3) Guided Repeated Reading, which relies exclusively on frequent reading to strengthen fluency skills. Similar to previous research, while participants in each intervention type made gains in their basic literacy skills, as stated by the authors, "all instructional programs . . . were relatively effective in helping students to improve their basic reading skill scores at posttest in comparison to pretest" (Sabatini et al 2011:127). Gains in literacy varied depending on specific skills evaluated, such that the largest gains were found in decoding skills ($d = .46$) and more modest gains for other skills such as word identification, reading comprehension, and sight word efficacy ($d = .19-.21$).

Overall, the findings from these recent efficacy studies suggest that struggling adult readers do often make modest to moderate gains in their literacy skills across varied types of reading interventions. However, a specific literacy intervention that seems to surpass other types in overall gains has not been identified to date. An alternative consideration is whether there are specific individual factors versus program variables that predict higher gains and skill retention.

Demographic variables, such as age, gender, socio-economic status, first language) may directly or indirectly be associated with success in literacy programs (e.g. being a non-native language learner in the Alamprese et al [2011] study). However, specific cognitive or linguistic skills contribute importantly to literacy skills, as well as to gains and/or maintenance of new skills. For example, Taylor, Greenberg, Laures-Gore, and Wise (2011) examined the relations among oral language skills, written language skills and reading comprehension in struggling adult readers. Specifically, they assessed how syntactic abilities (e.g. the rules and patterns of grammatical sentences and phrases) relate to reading development, a relationship that has been previously established in children (Bentin, Deutsch, & Liberman 1990, Gillon & Dodd 1995, Mokhtari & Thompson 2006, Nation & Snowling 2000). Knowledge of the rules and patterns of grammar did not predict level of reading comprehension; however, other oral language skills (phonology, expressive and receptive vocabulary) did. These findings by Taylor et al. (2011) identify existing language skills that are associated with better literacy in adults who are not necessarily involved in a literacy program.

In a study of component skill associated with better literacy, Strucker, Yamamoto & Kirsch (2007) reported that high levels of proficiency in word reading and vocabulary abilities represented "tipping points," or thresholds that, once reached, can support a reader's entry into moderate (beyond early high school level, or what IALS termed Level 3) literacy. The authors acknowledged that this level of literacy could change individuals' lives dramatically, particularly if given a burst of intensive, tightly focused instruction. However, the process or component skills involved in achieving these threshold skills were not explicated.

A cognitive process that has been identified as an important mediator of the acquisition of reading skills is 'automaticity', or the process by which words are recognized automatically, so that attentional capacity can be directed toward resources required for comprehension (Greenberg et al 2011). Laberge & Samuels (1974) proposed a model demonstrating that any complex skill, such as reading, consists of a variety of component processes, and when executing these skills, each component may require attention or cognitive effort. If this is the case, then such complex tasks would be extremely difficult to execute, as the total demand of content plus processes surpasses the available attentional resources, resulting in the mental effort of the activity being too great. As a result, less content can be processed and understood, and in the case of reading, comprehension abilities would be negatively impacted. However, if some of the component steps became 'automatized' or 'automatic', then fewer attention resources overall would be required to execute these complex skills. For a skilled reader, less attention is required for each component, by presumably having automatized each of the sub-skills (e.g. labelling, syntax and pronunciation/articulation, verbal reasoning) and even more importantly, having made their integration automatic as well.

Subsequent researchers have examined the key role of automaticity in the development and acquisition of reading skills, in a variety of populations. For example, Rasinski, Rikli, and Johnston (2009) explored the development of reading fluency in samples of children in the third, fifth, and seventh grades. Having previously established that reading rate (word recognition automaticity) is significantly correlated with reading comprehension in ninth graders (Rasinski, Padak, McKeon, Krug-Wilfong, Friedauer, & Heim 2005), Rasinski and colleagues (2009) sought to measure the relationship between prosody, another aspect of reading fluency, and comprehension. The results of this study revealed that at all three grade levels students who demonstrated greater prosody (i.e., phrasing, intonation, pace, etc) had higher levels of comprehension. This study demonstrates, along with a growing body of research, that increased reading fluency, whether measured by prosody or automaticity, can lead to improved

reading comprehension (Griffith & Rasinski 2004, Martinez, Roser, & Strecker 1999, Biggs, Homan, Dedrick, Rasinski & Minick, 2008).

Similarly, Bebko, Bell, Metcalfe-Haggert and McKinnon (1998a) explored the relationship between automaticity and language proficiency in a sample of profoundly deaf children. Results showed that automaticity, as measured by a rapid automatized naming (RAN) task, predicted a deaf child's effective use of language-based strategies, such as spontaneous rehearsal (Bebko, et al., 1998a). These findings support the assumption that ' . . . a child whose language skills are not as yet overlearned or automatized must expend additional processing resources using a language-based strategy compared to a child who has automatized these skills' (p.64). Bebko (1998b) further extended the important role of automatization in explaining the literacy difficulties experienced by deaf adolescents and adults.

Winn, Skinner, Oliver, Hale, and Ziegler (2006) reported on two instructional approaches compared to a control condition. In the listening while reading (LWR) condition, participants silently read a passage while listening to the passage being read aloud; the repeated reading (RR) condition only required the participant to silently read the passage. In a sample of 12 struggling adult readers, greater fluency resulted in both the LWR and RR conditions, as reflected by speed and accuracy gains (i.e. more words correct per minute). This association between quicker and more accurate processing and fluent reading, regardless of the approach used, suggests a need for additional research with adults focused on individual variables, which was the focus of the present project.

Using an experimental, prospective, longitudinal design, the present study examined the degree to which automatization was related to outcome and subsequent retention of skills following participation in general literacy programs. Reading level was operationalized by performance on a standardized measure, and change in that performance between the beginning and end of the program was measured. Level of reading 'automaticity' was operationalized by performance on a battery of cognitive information processing measures. A one-year follow-up of learners was completed in order to measure retention of reading skills following their involvement in general adult literacy programs. It was hypothesized that the degree of automatization of reading skills at the end of adult literacy program participation would be associated with the most successful levels of achievement, and with subsequent retention of skills, months later at follow-up testing.

Method

Participants

Adult participants with at least a grade equivalency of 3 or higher were recruited from programs across the Greater Toronto Area with the assistance of the Ontario Literacy Coalition. Researchers went to each classroom, explained the study and invited learners to participate in the initial screening process. One hundred and eight participants were screened, from which a sample of 47 people was chosen to continue on with the study (based on having English as their first/primary language). Of the 47 who met these criteria, 6 scored over the Grade 11 level during the first testing session. Due to this initially advanced level of literacy, their data are excluded from the present analysis due to potential ceiling effects. Thus, for Session 1 the sample consisted of 41 people, whose tested reading levels were between Grade 3 and Grade 11 equivalencies on the Canadian Adult Achievement Test (CAAT, Psychological Corporation 1988).

Over the next 6 months, substantial attrition occurred due to learners leaving the various programs and losing contact with the research program. At the second testing session, or the third month of their involvement with the program, there were 30 participants, and by the third session, at 6 months, there were 21. The follow-up session consisted of 20 participants and took place several months after they left the program. The attrition did not appear to be systematic, in that it did not affect the overall demographic profile of the sample in serious ways (see Table 1 for a summary of the demographics for each testing session).

	Testing 1	Testing 2	Testing 3	Follow-up
n	41	30	21	20
Mean Age in years (range)	37.02 (18-64)	36.85 (18-64)	37.94 (19-64)	38.45 (19-64)
Gender (female:male)	32:9	23:7	17:4	16:4
Mean Highest Grade Level of Education	9.58 (3-12)	9.50 (3-12)	9.38 (3-12)	9.25 (3-12)

Table 1 Demographic characteristics for participants at testing sessions 1, 2, 3 and follow-up

Measures and Materials

Reading Achievement Measures.

The Canadian Adult Achievement Test (CAAT, The Psychological Corporation 1988, 1995) was used as both a screening measure to establish eligibility for inclusion in the study, as well as an indication of level of mastery of reading. The CAAT is designed to determine present educational level and includes a battery of norm-referenced, multiple-choice tests, which are separated into a variety of levels. Prior to the first testing

session, the CAAT Select-A-Level screener was administered. This is a multiple choice test with 45 questions to determine an approximate skill level in the areas of literacy and numeracy and to determine which comprehensive CAAT assessment version (A, B, or C) would be most appropriate.

For the purposes of the current study, the 15 numeracy questions on the screener were not scored, so people received a score out of 30. An adjusted cutoff was used to identify the version of the CAAT to be administered, maintaining the same percent correct used when all 45 questions of the Select-A-Level are included. Therefore, participants who scored 21 or more out of 30 were administered the CAAT-C (the most difficult version) during the three testing sessions. Participants who scored less than 21 were given CAAT-B for the three testing sessions (16 of the 41 participants). This ensured that the learners were assessed with measures suitable to their skill level, and would complete a test version that was challenging but not frustrating.

In each session, reading proficiency was evaluated using the Vocabulary and Reading Comprehension subtests of the CAAT. The Vocabulary subtest consists of 34 multiple-choice questions that assess the knowledge and understanding of common words. The Reading Comprehension subtest involves reading several short passages and answering a total of 50 multiple choice questions about the content and implications of the passages. Participants' overall reading grade levels were determined by averaging their grade levels on the 2 subtests. The CAAT has shown to have strong internal consistency and moderate to high content validity (The Psychological Corporation 1988).

A second, less standardized instrument was also used as a measure of more functional reading decoding skills. The REALM (Rapid Estimate of Adult Literacy in Medicine; Davis, Long, Jackson, Mayeaux, George, Murphy & Crouch 1993) uses common medical words or lay terms for common illnesses and parts of the body arranged in three lists of 22 words (full version). Participants read the lists as completely as possible and were scored based on accuracy indicated by the dictionary pronunciation. This test taps into underlying reading and decoding skills. In using the REALM it is assumed that if basic reading skills such as word recognition and decoding are weak, particularly in a context that is functionally relevant for everyday living, then difficulties with higher-level literacy skills such as comprehension, are likely. Therefore, inability to decode simpler words on the REALM suggests low comprehension and may be used to identify low-level readers.

Information Processing Test Battery

This battery of three tests was designed to provide multiple measures of automaticity of underlying processes involved in reading.

Rapid Automatized Naming task (RAN)

The Rapid Automatized Naming (RAN) task was used to measure the automatization of access to lexical information and word recognition. Given that naming speed reflects the degree to which all levels of visual to semantic decoding have been automatized, performance on these types of measures has been found to be one of the primary factors associated with reading skill (e.g. Lovett 1987).

Participants were administered three versions of the RAN: 1) five words (watch, umbrella, key, scissors, comb) were listed randomly, in five rows, ten words per row, on a large sheet of paper; 2) five pictures were presented (watch, umbrella, key, scissors, comb) in a similar design as the first RAN version; and 3) participants were given five words equivalent to a grade level that was one below their current level of reading. For each task, participants were asked to read as many words as possible as quickly as they could within 30 seconds. Order of presentation of the versions was randomized across participants and scores for each version of the RAN were calculated based on the number of words named within the time limit.

The Stroop Test

The Stroop test (1992/1935) was used to provide an indication of automatization of reading and the need to inhibit the automatized response (Cox, Chee, Chase, Baumgardner, Schuerholz, Reader, Mohr, & Denckla 1997). Participants were first presented with names of colours (red, blue, green) printed in 5 columns in black ink on an 8.5 inch by 11 inch sheet of paper. Participants were asked to first read down the columns of words as quickly as possible. In the second task, participants were presented with Xs that were printed in the colours represented by the words in the first task. Lastly, participants completed a third task in which the words from the first task were reprinted in coloured ink that was incongruent with the word. For this last task, participants were required to name the colour of the ink that each word was printed in, and not read the word. Participants were given a maximum of 45 seconds to complete each of the three tasks.

The Stroop task produced two scores: *interference* scores were calculated for each participant with the following formula: (Stroop colour – Stroop colour word) / Stroop colour. Therefore, the larger the score (i.e., the more interference), the more automatized their reading response. In an effort to compare their performance to how they were *expected* to perform, *cognitive flexibility* scores were also calculated for each participant by subtracting predicted interference scores from actual interference scores.

The Dual Task

A common means of determining the degree of automatization of a primary task is to examine its impact on a secondary task that is performed concurrently. In the present study, the impact of the primary task of reading was measured by looking at the rate of decline in the performance of a secondary task, rapid tapping on a finger tapper. The underlying assumption was that the more automatized the reading is, the less effort and attention it requires. Interference should then be minimal on one's ability to perform a concurrent, automatized task such as finger tapping. Dual task interference has been noted as one of the most important operational indicators for automatization (Heuer 1996).

Three tapping tasks were administered. Participants were instructed to keep their hand flat on the tapper and use only their index finger for tapping. The first task consisted of participants simply tapping as quickly as possible for 15 seconds. For the second task, participants were asked to read aloud a passage selected from the Diagnostic Reading Scales (Spache 1981) equivalent to a Grade 2 reading level, while simultaneously tapping for 15 seconds. Lastly, participants were asked to read aloud a passage from the Diagnostic Reading Scales equivalent to a reading level one below their current grade level (based on CAAT scores), again while simultaneously tapping for 15 seconds. A reading level of one below their current grade level was chosen to provide some challenge, yet still be within their abilities. The Diagnostic Reading Scales ceiling at Grade 7, so those participants requiring material equivalent to a Grade 8 level or higher were provided with corresponding passages from the Wechsler Individual Achievement Test (Wechsler 2005). Administration of the latter two tasks was counterbalanced between participants. Two finger tapping scores were calculated for each participant: Dual task Low (Grade 2 level) and Dual task High (Grade level – 1 year). These tapping scores were divided by performance when tapping alone to reflect dual task cost – the decrease in performance when the primary finger-tapping task was performed concurrently with the secondary reading task. Lower decrement scores reflect greater automaticity.

Procedure

One hundred and eight participants were initially screened and general demographic data were collected (e.g. gender, first language, most used language at home and at workplace, previous education, reason for enrolling in the program, amount of time spent in the program to date, and the length of time they planned to be in the literacy program). Participants who had English as their first language or had been using English as their primary language for at least 5 years were asked to continue in the study. At

this time, the Canadian Adult Achievement Test (CAAT) Select-a-Level screener was also administered in order to determine which level of the CAAT each participant should complete for the first testing session.

Forty-one participants began the first testing session shortly after entry into the program. Upon entry into the study, demographics were gathered as part of the first session, such as general socio-economic level, and any identified learning challenges or medical conditions. Persons with identified learning disabilities were excluded from data analyses.

In order to observe changes in their skills over time, participants were tested periodically with the two measures of reading achievement, along with the information processing test battery designed to determine the degree to which underlying processes involved in reading are automatized. Test sessions occurred every 3 months until the participant left the program, or until 9 months passed, whichever occurred first. (Because the end point of a program is sometimes only determined by the individual's continued absence from the program, so is not always predictable, the regular testing provided an estimate of the individual's level of skill development near the end of their program, even if departure was unanticipated). A follow-up session with the same measures was also conducted after participants had been away from the learning programs for several months (ranging from 2-8 months).

Results

Reading level achievement across time is summarized in Table 2. At the first testing session, the mean reading grade level for the 41 participants, as determined by the Vocabulary and Reading Comprehension subtests of the CAAT, was 6.94 (range of 3.50-10.20). For the 30 participants who were still involved in the second testing session, the mean reading grade level was 7.33 (range = 2.80 – 11.50). By the third testing session, there were 21 participants and their reading level was 8.01 (range of 3.50 – 12.60). For participants who were present for both testing sessions, grade level scores are significantly higher from time 1 to time 2 [$t(29) = -2.337, p = .027$] and from time 2 to time 3 [$t(18) = -2.277, p = .035$]. Note that when comparing times 2 and 3, the n is only 19 because 2 people who participated in the third session had missed the second testing. For the 20 participants tested at the follow-up session after leaving the program, their mean reading level was 7.90 (range 4.70-12.60).

For the REALM, scores at Testing 1 had a mean of 43.95 words (range of 8 - 64). At testing 2 there was a mean REALM scores of 42.83 (range = 4 - 66), and at Testing 3 the mean was 41.43 (range of 6 - 64). In contrast to the CAAT scores, these REALM scores did not change

significantly between Time 1 and Time 2 [$t(29)$ = .249, p = .805] or Time 2 and Time 3 [$t(18)$ = -.585, p = .566].

Initial reading grade levels of the participants at Time 1 were not significantly correlated with any of the demographic variables (age, gender, educational level, amount of time spent previously in the program, or having English as a first language), or the changes in CAAT scores from testing 1 to 2, or 2 to 3. The initial reading grade level of the participants at Time 1 was, however, correlated with initial performance at that time on RAN 2, $r(40)$ =.408, p =.008, and RAN3, $r(40)$ =.474, p =.002. RAN1 was correlated, but less strongly, $r(40)$ =.272, p =.086.. The performance of the participants across the two literacy measures, CAAT and REALM, was strongly and significantly correlated [$r(40)$ = .67, p <.001] at each of the 4 testing sessions.

	Testing 1	Testing 2	Testing 3	Follow-up
n	41	30	21	20
Mean Reading Grade Level *(Standard Deviation)*	6.94 (1.91)	7.33* *(2.25)*	8.01* *(2.42)*	7.90 *(2.29)*
Reading Grade Level Range	3.50 – 10.20	2.80-11.50	3.50-12.60	4.70-12.60

* significantly greater than previous mean reading grade level (p < .05)

Table 2 Reading Grade Level for participants at testing sessions 1, 2, 3 and follow-up

The 30 people who continued with the study at Time 2 were compared to the 11 people who left the study after testing 1 to determine if there were any characteristics that could account for who left and who remained. However, age, gender, educational level, amount of time spent in their literacy programs prior to the start of the study, as well as scores on the CAAT and REALM at Time 1 were not significantly different between the groups (age, p = .546; gender, p = .732; educational level, p= .699; prior time in program, p = .11; CAAT, p = .430; REALM, p = .721).

Correlation analyses were conducted to determine which variables were associated with level of literacy change on the CAAT from Time 1 to Time 2. Results indicated that the combined score on the three RAN variables at Time 1 was borderline significantly correlated, p< .07, n = 30.

Regression analyses showed that none of the demographic variables (age, reading level, length of time spent in the program, etc.) were associated

with literacy improvement. In terms of the information processing measures, there were no significant predictors of level of literacy change from Time 2 to Time 3. However, level of literacy change from Time 1 to Time 3 was significantly associated with Dual Task High (performance on the dual task while reading a passage near grade level) at Time 1 ($r(21) = -.521$, $p = .015$).

Retention of gains in literacy skills as measured on the CAAT was defined by Grade level at Time 4 minus Grade level at previous testing divided by time away (in months). Of note, as aforementioned, mean reading level at Time 3 (8.01, range 3.50-12.60) was *not* significantly different from follow up at Time 4 (7.90, range 4.70-12.60). The differences were divided by time away in order to give the best estimate of the effects of time decay. Regression analyses indicated that the Stroop Interference variable was the strongest predictor of retention, approaching statistical significance, $F(1,18) = 3.177$, $p = .092$ $r = .387$; the corresponding r^2 indicates that Stroop Interference accounted for 15% of the variance in retention. This indicates that as reading became more automatized (thereby causing greater interference on the Stroop Task), the greater was the probability of retention. No other variables were strongly associated with retention.

Discussion

A number of studies have compared different approaches to literacy training in struggling adult readers, and these have found the majority of approaches examined to be effective in increasing literacy skills compared to no training, but little difference among different approaches (e.g., Greenberg, et al 2011, Sabatini, et al 2011, Winn, et al 2006). That focus on the relative efficacy of various programs has been complemented by a search for individual factors that are associated with higher gains in programs, and retention of those gains.

In the present study, we followed the longitudinal progress of learners, across approximately 6 months of program involvement, plus follow-up approximately a year from the beginning of the study. The prime focus was on the individual factors that may be predictive of retention of gains versus differences among program types. In contrast to a number of efficacy studies (Alamprese, et al 2011, Friedlander & Martinson, 1996, Greenberg, et al 2011, Mikulecky & Lloyd, 1997, Sheenan-Holt & Smith 2000) where small gains were seen among participants, our results were similar to Sabatini and colleagues' (2011) in that participants showed solid and significant gains in their literacy skills during participation in their varying programs. Participants in the present study showed a mean gain of one grade level in vocabulary and general reading comprehension in the 6 months of participation, and the gains were maintained at follow-up, for those

available for testing. No demographic or other pre-study variables were associated with outcome or with early departure from programs.

Our specific focus for this study was to determine cognitive information processing characteristics of participants that were associated with outcome and retention of gains. Automatization of component skills has been hypothesized to be associated with literacy acquisition in both children and adults. Consistent with this hypothesis, the RAN tasks, which are measures of automatized and rapid identification of words and pictures, were strong predictors of gains in reading grade level from testing sessions 1 to 2. In addition, regression analyses revealed that the strongest predictor of retention (approximately six months after leaving the programs) of gained literacy skills was participants' last previous Interference score on the Stroop task prior to follow-up testing. The assumption in Stroop testing is that the greater the interference that is experienced on the Stroop task, the more 'automatic' the person's reading was. This automaticity of reading was, in turn, associated with greater retention of the literacy skills learned in the program, with automaticity accounting for 15% of the total variance associated with retention. This is a strong and important finding, particularly given the small number of participants in the follow-up session ($n = 20$).

The results of the current project show that literacy improvements and retention of gains are associated with various reading automaticity measures, and these results may provide literacy practitioners with important information about how to promote the retention of reading skills for adult learners. It appears that simply achieving a grade level or level of improvement is not sufficient. To ensure retention of skills, those skills must become automatized, otherwise the complex act of reading will continue to be too difficult a task and learners will not persist with it. For all cognitive activities, there are limited resources that can be kept activated at the same time for ongoing processing of the environment. Prior to achieving automaticity levels, the mental effort required for reading is considerable, as component processes, such as decoding and organizing the decoded sounds into meaningful words, compete with efforts at comprehension of the materials for the available resources. Once the component processes have been automatized, more of the available resources can be dedicated towards comprehension, correspondingly increasing the fluency of reading. The role of 'over learning' or automaticity on the retention of skills is clearly an avenue for future research.

Several important implications for programs derive from these findings. If automatization of component reading skills were measured at the outset of a program, then it may be optimal for those with more automatized skills to have the focus of the program on increasing fluency

skills. An example of this approach would be a Guided Repeated Reading approach (Shore [2003] cited in Sabatini, et al 2011). However, the present findings imply that the target for this approach should be a high degree of speed and automaticity, reflected by measures such as performance on a Stroop task. Interestingly, more automaticity of reading is reflected by greater interference on the Stroop task, indicating that participants cannot suppress the salience of reading the text on the measure, even though the instructions are to ignore the printed words and report the colors in which they are printed. The strength of interference is assumed to result from the degree to which the participants are unable to suppress the reading response.

At the same time, a program pretest that identifies participants with poor automatization skills at outset, such as low speed of responding on a Rapid Automatized Naming task or other simple measures, would imply a different focus for instruction. The focus for these participants should likely be more towards a Corrective Reading approach (Engelmann, Carnine & Johnson 1999) adapted for adults, or similar approach, with a focus on initial grapheme-phoneme associations and on word recognition. Some improvement has been found in such a program with adults (Sabatini et al 2011), but not differentially compared with other program approaches.

For instructors who do not have the resources for pre- and post-testing of participants, it would be beneficial at least to estimate the degree of automatization of reading skills even informally. Some examples might include: having the same passages read aloud privately to the instructor early in a program and again later and comparing the time taken; or having learners compete an informal RAN-type task or Stroop-type task early and late in a program. Simplified and non-standardized versions of the RAN and Stroop are readily available through Internet searches. Although these are informal, such within-person comparisons would provide useful information about an individual's progress in automatizing component reading skills. In terms of resources, this informal monitoring would, nonetheless, require one-to-one evaluation time.

It is possible that attempts to compare the relative efficacy of specific programs in the past have been confounded by the presence of struggling readers with both types of needs, those with initial rapid grapheme-phoneme recognition difficulties, and those with better developed automatization skills, but still in need of further fluency development. Random assignment to program types in studies such as those by Alamprese et al (2011) and Sabatini et al (2011), while advantageous for isolating program variables, may have contributed to this confounding versus assignment based on individual processing characteristics. Clearer efficacy studies may result if participants are separated out by the degree of

automatization of initial component skills at the outset of their involvement, and the corresponding targeting of type of support to the specific processing skills of the individuals.

One limitation of this study, and many other longitudinal studies, is the attrition of participants over time. Our final sample was just over 50% of our initial participants (21 of 41). Attrition analyses comparing characteristics of those who departed to those who remained indicated that attrition was essentially random, introducing no noticeable biases into the data analyses. Monetary incentives were used to encourage continued involvement in the study; however, the incentives were based on participation in the adjunct information processing testing, not ongoing attendance at the literacy programs. Although participation in the adjunct testing was based on those continuing in the literacy programs, it is unlikely that the incentives served to motivate program attendance. Some well-resourced studies have maintained higher levels of retention across nine months or more of study (e.g. see Comings & Soricone 2007 for a review), but various factors in the lives of program participants, some predictable, some not, contribute to attrition. The present study was fortunate in that attrition appeared to be random in terms of participant characteristics. Nonetheless, the retention of adult literacy participants remains an ongoing challenge, both in terms of instructional effectiveness and to ensure robustness of research findings.

Acknowledgements

References

Alamprese, Judith, MacArthur, Charles, Price, Cristofer and Knight, Deborah (2011) Effects of a Structured Decoding Curriculum on Adult Literacy Learners' Reading Development, *Journal of Research on Educational Effectiveness*, vol 4, no 2, pp 154-172.

Bebko, James, Bell, Michelle, Metcalfe-Haggert, Alisa and McKinnon, Elaine (1998a) Language Proficiency and the Prediction of

Spontaneous Rehearsal Use in Children who are Deaf, *Journal of Experimental Child Psychology,* vol 68, no 1, pp 51-69.

Bebko, James. (1998b) Learning, Language, Memory, and Reading: The Role of Language Automatization and its Impact on Complex Cognitive Activities, *Journal of Deaf Studies and Deaf Education,* vol 3, no 1, pp 4-14.

Bentin, Shlomo, Deutsch, Avital and Liberman, Isabelle (1990) Syntactic Competence and Reading Ability in Children, *Journal of Experimental Child Psychology,* vol 49, no 1, pp 147-172.

Biggs, Marie, Homan, Susan, Dedrick, Dedrick, Minick Vanessa and Rasinski, Timothy (2008), Using an Interactive Singing Software Program: A Comparative Study of Middle School Struggling Readers. *Reading Psychology*, vol 29, no 3, pp195-213.

Canadian Adult Achievement Test (1988) Harcourt Canada Ltd. and The Psychological Corporation, Toronto.

Cassidy, Jack, Valadez, Corinne Montalvo, Garrett, Sherrye Dee and Barrera, Estanislado (2010) Adolescent and Adult Literacy: What's hot, What's not, *The Journal of Adolescent and Adult Literacy*, vol 53, no 6, pp 448 - 456.

Cohen, J (1988) Statistical power analysis for behavioral sciences, (2nd ed.), Hillsdale, NJ: Lawrence Erlbaum.

Comings, John and Soricone, Lisa (2007) *Adult literacy research: opportunities and challenges, NCSALL Occasional Paper,* National Centre for the Study of Adult Learning and Literacy (NSCALL), Boston, MA.

Cox, Christiane, Chee, Elsabeth, Chase, Gary, Baumgardner, Thomas, Schuerholz, Linda, Reader, Mark, Mohr, Jennifer, and Denckla, Martha (1997) Reading Proficiency Affects the Construct Validity of the Stroop Test Interference Score, *The Clinical Neuropsychologist*, vol 11, no 2, pp 105-110.

Engelmann, Siegfried, Carnine, Linda and Johnson, Gary (1999) *Corrective Reading*, SRA McGraw-Hill, Columbus, OH.

Friedlander, Daniel and Martinson, Karin (1996). Effects of Basic Mandatory Education for Adult AFDC Recipients, *Educational Evaluation and Policy Analysis*, vol 18, no 4, pp 327-337.

Gillon, Gail and Dodd, Barbara (1995) The Effects of Training Phonological, Semantic, and Syntactic Processing Skills in Spoken Language on Reading Ability, *Language, Speech, and Hearing in the Schools,* vol 26, no 1, pp 58-68.

Gombert, Jean Emile (1994) How do Illiterate Adults React to Metalinguistic Training? *Annals of Dyslexia,* vol 44, no 1, pp 250 - 269.

Greenberg, Daphne, Wise, Justin, Morris, Robin, Fredrick, Laura, Nanda, Alice and Pae, Hye-K (2011) A Randomized Control Study of Instructional Approaches for Struggling Adult Readers, *Journal of Research on Educational Effectiveness*, vol 4, no 2, pp 101-117.

Griffith, Loraine and Rasinski, Timothy (2004) A Focus on Fluency: How One Teacher Incorporated Fluency with her Reading Curriculum. *The Reading Teacher*, vol 58, no 2, pp 126-137.

Heuer, Hegele (1996) Dual-task performance, in Neumann, Odmar and Frans Sanders, Andries, Eds *Handbook of Perception and Action, Volume 3: Attention* Academic Press, London, pp. 113-153.

Kozol, Jonathan (1985) *Illiterate America,* Anchor Doubleday, New York, NY.

Kutner, Mark, Greenberg, Elizabeth, Jin, Yin and Paulsen, Christine (2006) *The health literacy of America's adults: Results from the 2003 national assessment of adult literacy. (No. NCES 2006-483)*, U.S. Department of Education, National Center for Education Statistics, Washington, D.C.

LaBerge, David and Samuels, S. Jay (1974) Toward a Theory of Automatic Information Processing in Reading, *Cognitive Psychology*, vol 6, no 2, pp 293-323.

Lovett, Maureen (1987) A Developmental Approach to Reading Disability: Accuracy and speed criteria of normal and deficient reading skill, *Child Development*, vol 58, no 1, pp 234-260.

Martinez, Miriam, Roser, Nancy and Strecker, Susan (1998/1999) "I Never Thought I Could be a Star": A Reader's Theater Ticket to Fluency, *The Reading Teacher*, vol 52, no 4, pp 326-334.

Mikulecky, Larry and Lloyd, Paul (1997) Evaluation of Workplace Literacy Programs: A Profile of Effective Instructional Practices, *Journal of Literacy Research*, vol 29, no 4, pp 555-585.

Mokhtari, Koulder and Thompson, Brian, (2006) How Problems of Reading Fluency and Comprehension are Related to Difficulties in Syntactic Awareness Skills among Fifth Graders', *Reading Research & Instruction,* vol 46, no 1, pp 73-96.

Morais, Jose, Content, Alain, Bertelson, Paul, Cary, Luz and Kolinsky, Regine (1988) Is there a Critical Period for the Acquisition of Segmental Analysis? *Cognitive Neuropsychology,* vol 5, no 3, pp 347-375.

Nation, Kate and Snowling, Margaret (2000) Factors Influencing Syntactic Awareness Skills in Normal Readers and Poor Comprehenders, *Applied Psycholinguistics*, vol 21, no 2, pp 229-241.

National Adult Literacy Survey (1993) National Center for Education Statistics, Washington, DC.

National Research Council (1998) *Preventing Reading Difficulties in Young Children*, National Academy Press, Washington, DC.

Rasinski, Timothy, Padak, Nancy, McKeon, Christine, Wilfong, Lori, Friedauer, Julie and Heim, Patricia (2005) Is Reading Fluency a Key for Successful High School Reading? *Journal of Adolescent and Adult Literacy,* vol 49, no 1, pp 22-27.

Rasinski, Timothy, Rikli, Andrew and Johnston, Susan (2009) Reading Fluency: More than Automaticity? More than a Concern for the

Primary Grades? *Literacy Research & Instruction*, vol 48, no 4, pp 350-361.

Roman, Sarah (2004) Illiteracy and Older Adults: Individual and Societal Implications. *Educational Gerontology*, vol 30, no 2, pp 79-93.

Sabatini, John, Shore, Jane, Holtzman, Steven and Scarborough, Hollis (2011) Relative Effectiveness of Reading Intervention Programs for Adults with Low Literacy, *Journal of Research on Educational Effectiveness*, vol 4, no 2, pp 118-133.

Satherley, Paul, Elliot, Lawes and Sok, Shelia (2008) *The Adult Literacy and Life Skills (ALL) Survey: Overview and International Comparisons*, Comparative Education Research Unit, Research Division, Ministry of Education, Wellington, New Zealand.

Sheehan-Holt, Janet and Smith, M. Cecil (2000) Does Basic Skills Education Affect Adults' Literacy Proficiencies and Reading Practices? *Reading Research Quarterly*, vol 35, no 2, pp 226-243.

Spache, George D. (1981) *Spache Diagnostic Reading Scales*, CTB- McGraw-Hill, Monterey, CA.

Statistics Canada. (2003). Building on our Competencies: Canadian Results of the International Adult Literacy and Skills Survey (Catalogue number 89-617-XIE), from Statistics Canada, retrieved 7 Aug 7, 2011 from http://www.statcan.gc.ca/.

Sticht, Thomas and Armstrong, William (1997) *Adult Literacy In the United States: A Compendium of Quantitative Data*, retrieved 14 Nov 2012 from http://www.nald.ca/fulltext/adlitUS/intro.htm.

Stroop, John (1992) Studies of Interference in Serial Verbal Reactions, *Journal of Experimental Psychology General*, vol 121, no 1, pp 15-23. (Reprinted from *Journal of Experimental Psychology*, 1935, vol 18, no 6, pp 643-662).

Strucker, John, Yamamoto, Kentaro, and Kirsch, Irwin (2007) *The Relationship of Component Skills of Reading to IALS Performance: Tipping Points and Five Classes of Adult Literacy Learners* (NCSALL Reports #29), Harvard Graduate School of Education, National Centre for the Study of Adult Learning and Literacy Cambridge, MA.

Taylor, Nicole, Greenberg, Daphne, Laures-Gore, Jacqueline and Wise, Jusint (2011) Exploring the Syntactic Skills of Struggling Adult Readers. *Reading and Writing: An Interdisciplinary Journal*, vol 25, no 6, pp 1-18.

Torgerson, Carole, Porthouse, Jill and Brooks, Greg (2005) A Systematic Review and Meta-Analysis of Randomized Controlled Trials Evaluating Interventions in Adult Literacy and Numeracy, *Journal of Research in Reading*, vol 26, no 3, pp 234-255.

Vanderberg, Laura, Pierce, Margaret and Disney, Laurel (2011) Reading Intervention Outcomes for Adults with Disabilities in a Vocational Rehabilitation Setting: Results of a 3-year Research and

Demonstration Grant, *Rehabilitation Counseling Bulletin*, vol 54, no 4, pp 210-222.

Venezky, Richard and Sabatini, John (2002) Introduction to this Special Issue: Reading Development in Adults. *Scientific Studies of Reading*, vol 6, no 3, pp 217-220.

Walker, Maurice, Udy, Karl, Pole, Nicholas, May, Steven, Chamberlain, Glenn Stephen and Sturrock, Fiona (1997) *Adult Literacy in New Zealand: Results from the International Adult Literacy Survey*, Ministry of Education, New Zealand.

Weiss Barry, Francis, Laurie, Senf, Janet, Heist, Kim and Hargraves, Rie. (2006). Literacy Education as Treatment for Depression in Patients with Limited Literacy and Depression: A Randomized Controlled Trial. *Journal of General Internal Medicine*, vol 21, no 8, pp 823-828.

Wechsler, David. (2005). Wechsler Individual Achievement Test 2nd Edition (WIAT-II). London: The Psychological Corp.

Winn, Beth, Skinner, Christophexr, Oliver, Renee, Hale, Andrea and Ziegler Mary (2006) The Effects of Listening while Reading and Repeated Reading on the Reading Fluency of Adult Learners, *Journal of Adolescent and Adult Literacy,* vol 50, no 3, pp 196-205.

Woodcock, Richard and Johnson, Mary (1990) *Woodcock-Johnson Psychoeducational Battery-Revised*, DLM Teaching Resources, Allen, TX.