Applied Psycholinguistics **14** (1993), 369–385
Printed in the United States of America

Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition

NORMAN S. SEGALOWITZ
Concordia University

SIDNEY J. SEGALOWITZ
Brock University

ADDRESS FOR CORRESPONDENCE
Norman Segalowitz, Psychology Department, Concordia University, 7141 Sherbrooke Street West, Montreal, Quebec H4B 1R6, Canada

ABSTRACT
Practice on cognitive tasks, in general, and word recognition tasks, in particular, will usually lead to faster and more stable responding. We present an analysis of the relationship between observed reductions in performance latency and latency variability with respect to whether processing has merely become faster across the board or whether a qualitative change, such as automatization, has taken place. The coefficient of variability (CV) – the standard deviation of response time divided by the mean latency – is shown to be useful for this purpose. A cognitive interpretation of the CV is given that relates it to issues of skill development. Data from second language learners' word recognition performance and from a simple detection task are presented which confirm predictions drawn from this interpretation of the cognitive significance of the CV. Initial improvement in a second language word recognition task was interpreted as involving more efficient controlled processing, which later gave way to automatization. The implications of this index of skill are discussed in relation to second language development and the general issue of automaticity of processing components in cognitive skills.

A widely documented generalization about the effects of practice on skilled behavior is that it leads to faster and more stable performance latencies (for reviews, see Logan, 1988; Newell, 1991; Newell & Rosenbloom, 1981). There is also some evidence that this appears to be the case with word recognition skills as tapped by a lexical decision task. For example, we reanalyzed some data from a previously published study (Segalowitz & Hébert, 1990) involving bilinguals who read more quickly in their first

0142-7164/93 $5.00 + .00

language (L1) than in their second (L2) and who presumably had had a great deal more practice reading in their L1. We found that the bilinguals exhibited both longer lexical decision response times (RT) and greater RT variability in L2 than in L1. Even within L1, the faster responding subjects, who were presumably more skilled, tended to exhibit less variability in their responses. Thus, overall, the faster the readers were able to recognize an item as a word, the more stable their RT was.[1] While this association between RT and RT stability as a function of skill and practice may seem fairly obvious, the explanation for it is not; the search for an account of practice effects is a central goal of researchers concerned with understanding skill acquisition, in general (Ackerman, 1987; Fleishman & Mumford, 1989), and in reading skill, in particular (Perfetti, 1985; Segalowitz, 1991; Segalowitz, Poulsen, & Komoda, 1991; Stanovich, 1990).

One useful framework for understanding how practice improves performance is found in the literature on the distinction between automatic and controlled operation of the processes underlying a given cognitive activity (e.g., Neely, 1977, 1991; Posner & Snyder, 1975). While various competing possibilities have been advanced as to how best to characterize this distinction (Kahneman & Triesman, 1984), several points of consensus can be noted. The first is that most complex cognitive skills (e.g., reading, Perfetti, 1985; reading a second language, Segalowitz, 1986; playing a musical instrument, Sloboda, 1985) involve the use of processes that require cognitive effort. These fall under strategic control and require attention; in operational terms, they involve decision-making and are vulnerable to interference. Second, under appropriate conditions, practice leads to faster, less effortful, and more stable performance. That is, elements of the skill become automatized in the sense that some underlying components no longer fall under strategic control (decision-making has been reduced) and no longer require attentional effort (performance becomes more immune to interference from competing cognitive demands). Unfortunately, in any given situation, it is seldom a matter of whether performance is simply automatic or under strategic control. Virtually all complex processing will involve a blend of mechanisms, some of which operate automatically, while others are under strategic control (see Jacoby, 1991, for an especially insightful discussion). The researcher who is interested in studying skill development faces the challenge of devising research methods for identifying when a particular experience (e.g., extensive practice in reading) results in more automatized performance (e.g., more automatic word recognition; see Favreau & Segalowitz, 1983; Segalowitz, 1991; Segalowitz et al., 1991).

There are a variety of theoretical accounts of how automaticity may be brought about by practice. Schneider, Dumais, and Shiffrin (1984) argued that processes that are characterized as controlled at the outset become automatized (less effortful) through practice over time. Cheng (1985), on the other hand, suggested that practice leads to a structural reorganization of underlying components. Logan (1988) proposed that, through practice, the performer shifts from reliance on algorithmic procedures to direct memory look-up. A fourth account suggested that practice renders performance more "ballistic" and less subject to monitoring by specialized attentional mechanisms by eliminating or reducing reliance on higher level supervisory processes (Shallice, 1982). Stanovich (1990) also discussed a shift from reliance on mechanisms that depend on general knowledge to processing modules, or what Fodor (1983) referred to as "informationally encapsulated" processing units. In these units, information analysis takes place without reference to other, higher level information. An example might be word recognition that occurs without reference to knowledge of a sentence or story context. What is common to all these accounts is that controlled processing is seen to involve mechanisms whose execution times will be relatively slow and variable and whose influence on performance diminishes as performance becomes more automatized.

Thus, controlled processes are presumed to contribute significantly, both to the slowing of performance and to latency variability (Ackerman, 1987). Practice reduces or eliminates their influence and thereby improves (i.e., speeds up and stabilizes) performance. We make here a broad distinction between two types of change that may occur as a result of practice. The challenge before us is how to know whether, in a given case, observed changes in performance latency and latency variability reflect (a) *speed-up effects*, whereby processes function faster and more stably, but otherwise continue to operate in the same manner (with no gain in automaticity), or (b) *qualitative change*, whereby performance is faster and more stable because the contributions of relatively slow and variable controlled processing components have been eliminated or reduced (with a gain in automaticity).

Facilitory effects and speed-up

One way in which practice may benefit performance is through a general *facilitory effect*. This effect involves change that is essentially quantitative, corresponding to an across-the-board acceleration or speed-up (Anderson's [1982] "strengthening"; see also MacKay, 1982) of the processes involved in executing a given task. The effect of practice in this situation can be illustrated with the following hypothetical example. Consider, for example, Table 1A, which indicates 10 hypothetical component processes that form a blend of mechanisms underlying some complex activity, such as word recognition, in a lexical decision task. Since most processes are likely to operate at least partly in parallel and overlapping fashion (Townsend, 1974), for the sake of simplicity, we will assume that only the *rate-determining* aspects of these underlying processing mechanisms need to be considered here. These may be assumed to contribute additively to RT and variability. The latency column of the table shows hypothetical execution times for the 10 components. Associated with each component is an indication of the variability of their execution times (the Var. column). The corresponding standard deviations (*SD*) are shown to facilitate the calculation of the coefficient of variability (CV). Five of these hypothetical processes (components 1-5) are shown in the table as relatively fast acting

Table 1. *Illustrative hypothetical example of quantitative and qualitative effects of practice on latency, standard deviation (*SD*), and the coefficient of variability (CV)*

Component	Latency (arbitrary time units)	*SD*	Var.	CV
A. With ten hypothetical rate-determining components underlying task execution				
1	10	2.0	4.00	.200
2	34	6.8	46.24	.200
3	44	8.8	77.44	.200
4	54	10.8	116.64	.200
5	58	11.6	134.56	.200
6	64	25.6	655.36	.400
7	80	32.0	1024.00	.400
8	92	36.8	1354.24	.400
9	94	37.6	1413.76	.400
10	98	39.2	1536.64	.400
Overall	**628**	**79.77**	6362.88	**.127**
B. With practice, speeding all components by a factor of 0.5				
1	5	1.0	1.00	.200
2	17	3.4	11.56	.200
3	22	4.4	19.36	.200
4	27	5.4	29.16	.200
5	29	5.8	33.64	.200
6	32	12.8	163.64	.400
7	40	16.0	256.00	.400
8	46	18.4	338.56	.400
9	47	18.8	353.44	.400
10	49	19.6	384.16	.400
Overall	**314**	**39.88**	1590.52	**.127**
C. With practice, eliminating the 5 slowest components				
1	10	2.0	4.00	.200
2	34	6.8	46.24	.200
3	44	8.8	77.44	.200
4	54	10.8	116.64	.200
5	58	11.6	134.56	.200
Overall	**200**	**19.46**	378.88	**.097**

Note: SD = standard deviation of latency; Var. = variability of latency; CV = *SD*/latency.

and stable, corresponding perhaps to more automatized pattern-analyzing mechanisms. The remaining 5 (components 6–10) are relatively slow and variable, corresponding perhaps to controlled, effortful decision-making components. The overall performance RT will be the sum of the processing latencies for the underlying rate-determining components. The *SD* of the overall performance RT will be the square root of the sum of the variances for the latencies of the underlying components.

Table 1B indicates what happens if skill progresses to the point where all components operate 50% faster, but where, otherwise, there is no qualitative change in the way the cognitive act is carried out. The table shows that the latency and *SD* for each component have been reduced by 50%, and that both the overall RT and *SD* are reduced by 50%. This reduction in the overall variability is, of course, simply an arithmetical consequence of the speed-up of each component. Such reduction in variability could not be cited as evidence of increased automaticity of processing since, other than the speed-up, there has been no change in the way in which the underlying processes behave.

The coefficient of variability

How can we distinguish between (a) a situation in which both latency and variability have been reduced because of speed-up effects, and (b) a situation in which there has been a qualitative change due to the dropping out of some processing components or the modularization of processing? What is needed here is an index that combines information about latency and variability in such a way that we can see that the situation shown in Table 1A is not qualitatively different from the one shown in 1B. A useful measure for this purpose is CV, defined as *SD* divided by mean RT. This measure indicates the variability *for a given level of response latency* (see, e.g., Guilford, 1942). As Table 1A and 1B shows, the CV remains unchanged, even though latency and *SD* dropped in value in the latter compared to the former.

The CV is an index that up to now has not been used widely in cognitive psychological research, but it has been used in a number of other areas of psychology, ranging from studies of variability in components of fixed action patterns in animal behavior to the study of RT in brain-injured patients (Bekoff, 1977; Carlton, Robertson, Carlton, & Newell, 1985; Dawkins & Dawkins, 1973; McManus, Kemp, & Grant, 1986; Rastatter & Blair, 1984; Segalowitz, Unsal, & Dywan, 1992; Slater, 1978; Watson & Livesey, 1982; Weismer & Elbert, 1982). In all these cases, the CV has provided researchers with a measure of variability that is corrected for the latency of responding. In these examples, however, no theoretical significance has been attached to the CV. As we shall see, it is possible to interpret properties of the CV as indicating changes other than simple speed-up effects.

Restructuring effects and qualitative change

Practice can also lead to performance gains through qualitative changes in the functioning of the underlying processes through a restructuring effect.

Thus, instead of simply speeding up the operation of the component processes, practice may set conditions for them to become organized differently; for selected, inefficient processes to drop out; for new, more efficient processes to replace older, less efficient ones; or for some mixture of all these possibilities to occur (e.g., Crossman, 1959). In all these cases, the qualitative change is brought about by the elimination of some of the processes whose variability-to-latency ratio is relatively high. Table 1C indicates what will happen in our hypothetical illustration if skill has progressed to the point at which the five high CV rate-determining components (components 6–10) have been eliminated from execution of the task. The overall RT will be shorter and the *SD* will be reduced, as was the case in the speed-up example discussed earlier (Table 1B). However, this time we can see that the CV also drops (from .127 to .097), whereas in the speed-up example, it remained unchanged. This is due, of course, to the removal of the elements with high CV from the summation of the variances. In general, if performance becomes faster and more stable because one or more high CV components have dropped out, then the overall CV will also be reduced, along with the overall RT. In such a case, the CV will correlate with RT, since reductions in the overall RT will be accompanied by reductions in CV.

To summarize, practice may improve skill by reducing reliance on the slower, less stable, high CV components: those that we typically imagine to involve decision-making and supervisory functions. The improvement in this case will not be due simply to an across-the-board reduction in speed of operation, but rather to a *change in the number and relative variability of the components* contributing to the overall variability. Changes in *SD* will not be linearly proportional to the changes in the overall RT, since a qualitative change will have taken place. In fact, the proportional reduction in variability will be greater than the overall proportional reduction in latency. That is, CV will be positively correlated with overall RT, since every reduction in the RT will be accompanied by a more-than-proportional reduction in variability. Thus, if the RT is reduced with training and changes in CV correlate significantly with changes in the RT, then we may conclude that some kind of restructuring or reorganization has occurred to reduce the *SD* over and above what would be expected by virtue of the RT becoming faster.

Phases of skill development

One can relate this analysis to different phases of skill development. Initially, one can expect training to lead to a general facilitory phase, in which practice over a number of training sessions has the effect of facilitating performance in a global manner. Component processes become faster across the board, but there is no structural change in the organization of underlying mechanisms. The change reflects a speed-up effect only: across sessions, RT and variability are reduced, while the CV remains stable. The correlation between CV and RT does not differ significantly from zero, either across individuals in this phase or within a given individual across performances.

In a subsequent automatization phase, practice leads to structural changes. There is a significant reduction of reliance on the slower, less efficient, and controlled supervisory mechanisms. The change here is qualitative: not only the RT and variability, but also the CV, are reduced, since the whole processing system is becoming relatively less variable in relation to the overall RT, due to the elimination or replacement of slower, less stable components. The CV and RT will ~~not~~ correlate positively. Finally, in a third asymptotic phase, performance reaches optimal levels. Very little RT speed-up is possible (due to floor effects).

In the current study, we tested the hypothesized relations between RT, *SD*, and CV by examining cross-subject variation in performance on two tasks. The first was a simple RT task aimed at investigating the relationship between RT, *SD*, and CV in which no qualitative changes were to be expected. Here we expected a priori that individual differences in simple RT would reflect only the speed of constituent components and not qualitative differences in skilled performance. The second was a lexical decision task in English, the subjects being French speakers enrolled in an English language summer program. The lexical decision task also included a repetition condition permitting a within-subjects analysis. Because this lexical decision task required more complex cognitive activity than the simple detection task, we expected individual differences in lexical decision RTs to be related to qualitative differences in skilled performance.

EXPERIMENT 1

The goal of this first experiment was to investigate the relationship between RT, *SD*, and CV for a simple task in which individual differences in performance are *not* likely to involve differential use of effortful processes. A simple RT task requiring the detection of the onset of a visual stimulus was used. It was assumed that for this simple task, individual differences would not arise from the differential use of effortful processes, but rather primarily from differential speeds of the underlying component processes. In terms of the illustrative hypothetical example given earlier, differences between individuals should be modeled after the speed-up differences that characterize Table 1A and 1B. Thus, according to this theoretical analysis, it was expected that *SD* would correlate significantly with mean RT across subjects, but that CV would not.

Method

Subjects. Sixty-six French speaking students enrolled in a summer program in English as a second language (ESL) at Brock University participated in this experiment. They ranged in age from 18 to 25 years.

Procedure. The subjects responded to the onset of a computer-controlled visual stimulus (a square) appearing after a random ISI varying from 1000 to 3000 msec by pressing a response key. Simple RTs were collected for 30 trials.

Results and discussion

For these analyses, all RTs greater than 1000 msec were discarded as misses and considered indicative of attention lapses (these amounted to less than 2% of all trials). The mean RT for this task was 264 msec, the mean *SD* was 54 msec, and the mean CV was .200. More important for our purposes, *SD* correlated significantly with RT ($r = .61, p < .001$), while CV did not, $r = .19$, $t(64) = 1.55, p < .20$ (see Figure 1). The difference between these correlations was significant, $t(61) = 25.5, p < .001$. This supports the theoretical analysis presented earlier. The data show that individual differences in *SD* for a simple detection task do not necessarily reflect a qualitative processing difference between subjects. Rather, they are consistent with the idea that the faster subjects simply executed the component processes underlying the task more quickly. Their reduced variability (*SD*) is simply a reflection of an increased speed of responding. This is because *SD* as a measure of performance confounds variability with absolute RT. The CV, on the other hand, did not change as a function of RT. This is consistent with the idea that performance differences on such a simple task will reflect the speed of operation of the underlying processes and not their organization.

EXPERIMENT 2

The goal of this second, and main, experiment was to investigate the relationship between RT, *SD*, and CV in a task in which individual differences are more likely to involve differential use of effortful processes. For this we used a speeded (timed) lexical decision task as a measure of word recognition (den Heyer, Goring, Gorgichuk, Richards, & Landry, 1988), with English-language targets and with subjects whose skill in English as a second language varied. Making a lexical decision is clearly a more complex cognitive task than the simple detection task used in the previous experiment. It was expected that this task would involve differential reliance on effortful (attentional) processes, with less skilled subjects being more dependent on such processes than highly skilled subjects. In fact, Favreau and Segalowitz (1983) demonstrated that even otherwise-fluent L2 readers may exhibit less automatic word recognition in their L2 compared to their L1. Thus, it was expected that, unlike the results in Experiment 1, CV would correlate significantly with RT across individuals, since faster subjects would carry out word recognition with fewer effortful (i.e., slow and highly variable) components in their performance. In addition, a selected set of target items was repeated throughout the course of the experiment to allow us to examine within-subject relationships between RT, *SD*, and CV.

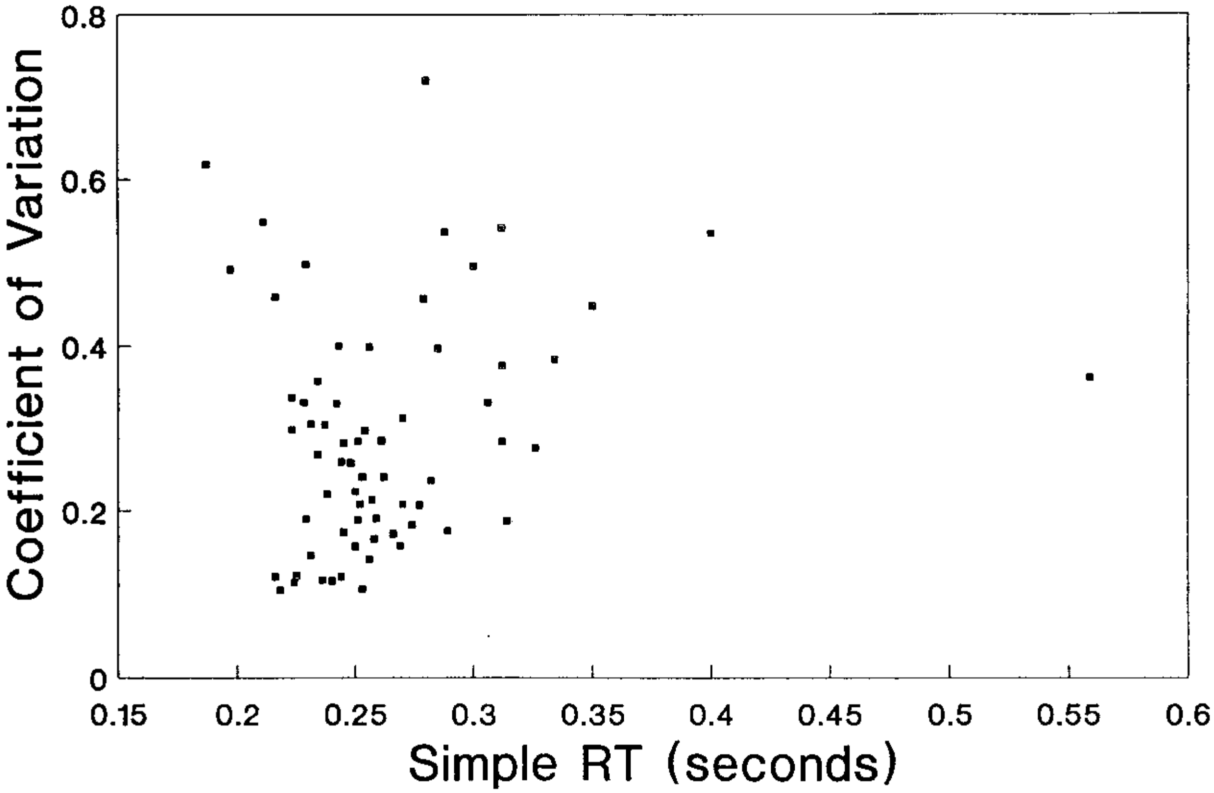

Figure 1. Plot of simple reaction times versus coefficient of variation for data from Experiment 1. (Note: $r = .194$, n.s. [$n = 66$])

Method

Subjects. The same 66 subjects participated as in Experiment 1. They ranged in English fluency from beginner to near fluent, as determined by global language skill placement tests administered by the English language program they were attending. Two subjects did not complete language background questionnaires, and so their data were dropped from these analyses, leaving 64 subjects.

Procedure and materials. The subjects performed a lexical decision task with 284 English words and nonwords. Each trial comprised the presentation of a target item (with a 3000 msec response deadline), followed by an intertrial interval of 1500 msec. The targets for the lexical decision task were constructed as follows: 35 items served as *baseline* words, 90 items (15 items repeated 6 times each) served as *repetition* words, 35 items served as *homophone* words (e.g., *pear*; these are not considered further here), and 124 items were *pronounceable* nonwords (e.g., *breek*). The first four trials used nonword targets.

Results and discussion

For purposes of this study, we used latency to respond to baseline words as an indicator of the level of word recognition skill. Other measures could have been used, such as performance on global tests of second language proficiency, tests of vocabulary knowledge, or tests of the speed of reading text. We selected response latency to baseline words, since this clearly reflected some aspects of word recognition without involving many other components of language skill that would necessarily be tapped by other measures (e.g., syntactic skills, depth of knowledge of word meaning, ability to integrate second language information across a text, etc.). Thus, on the basis of response latencies to baseline words, we labeled subjects with an overall relatively slow latency as being at the less skilled end of the word recognition skill continuum (for our sample) and subjects with an overall relatively fast latency as being at the skilled end of the word recognition skill continuum.

Primary analysis of the baseline word data. RTs greater or less than 2 *SD*s from the mean for a given subject, across all word conditions, were replaced with a value of 2 *SD*s from the mean. The mean RT, *SD*, and CV were computed for each subject across the 35 baseline word trials. Mean RT was 948 msec, mean *SD* was 324 msec, and mean CV was .323. As expected, the *SD* correlated significantly with RT ($r = .90$, $p < .001$), and CV correlated significantly with RT ($r = .72$, $p < .001$) (see Figure 2). Thus, faster subjects showed less variability than slower subjects – a difference that was more than the proportional reduction that would have simply reflected faster processing. This result is consistent with the expectation

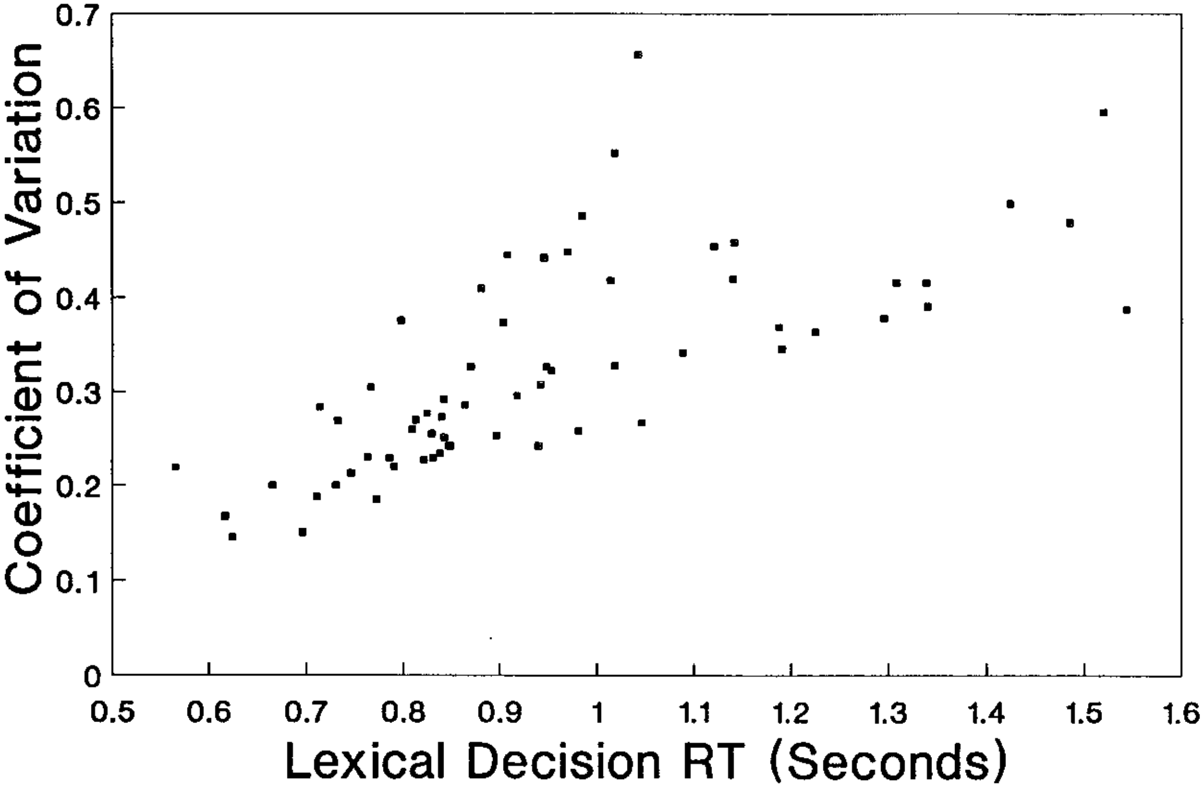

Figure 2. Plot of lexical decision reaction times versus coefficient of variation for data from Experiment 2. (Note: $r = .718$, $p < .00001$ n.s. [$n = 66$])

that, for this complex task, faster subjects used fewer effortful processes that were slower acting or highly variable.

Secondary analysis of the baseline word data. In this analysis we first examined the relation between RT, *SD,* and CV within the group of 22 subjects whose mean RTs located them at the extreme slow end of the word recognition skill continuum (the slowest one-third) and hence placed them closer to the initial phase of skill development. The subjects in this slow group were likely to differ from each other primarily in the speed with which many underlying components of word recognition were executed and less so in terms of the extent to which effortful processes played a role. A similar analysis was conducted on the data from the 22 subjects whose RTs placed them at the fast end of the continuum (the fastest one-third). Here, because these subjects may be expected to be in the intermediate and advanced stages of skill development, individual differences should begin to reflect a differential reliance on underlying effortful processes (Ackerman, 1987).

To test for these possibilities, separate correlational analyses were carried out on the groups comprising the slowest and fastest readers, the mean RT (and CV) for the slowest group being 1203 msec (.422) and for the fastest 745 msec (.233). The correlation of CV with mean RT for the 22 slowest performers was .20 (n.s.), while for the 22 fastest performers, the correlation was $r = .55$ ($p < .001$). Although this difference (between $r = .20$ and $r = .55$) did not reach statistical significance, when the fastest group was expanded to include the middle-range subjects, their correlation rose to .69 ($p < .001$), and the difference between the two groups was significant ($p < .05$). This provides further support for the idea that the relation of CV to RT reflects the degree to which underlying processing is automatic (i.e., not dependent on slower, more variable, component processes).

Analysis of the repetition word data. The purpose of this analysis was to investigate the relation of RT and CV to each other when practice effects are present within the task. Therefore, we analyzed the mean RTs and the correlations of CV with RT across the 64 subjects for the first presentation of each repetition word and for the last (sixth) presentation for each repetition word. Normally, lexical decision RT is faster with a repeated presentation of a target item than with the first presentation (den Heyer et al., 1988; Logan, 1990), and this was the case in our study: mean RT for first presentation = 880 msec; for last presentation = 750 msec; $t(63) = 5.48$, $p < .001$. Repetition of a target item should reduce the subject's need to rely on effortful, attentional processes involved in the early perceptual processing of the target. There should be gains in perceptual fluency with repetitions (Jacoby & Brooks, 1984), gains that are reflected in a positive correlation between CV and RT. Indeed, this is what was obtained. Although the correlation between CV and RT for responses to both the first presentation ($r = .427$) and the last presentation ($r = .666$) were significant ($n = 64$, $p < .001$), they differed from each other ($z = 1.92$, $p < .055$). According to the interpretation developed earlier, this increase in the correlation between CV and RT from the first to the last presentation reflects a gain in automaticity with practice.

We can expect, however, to find a difference in pattern from first to last presentation of repeated targets when we compare subjects who are relatively very skilled at word recognition (those at the fast end of the skill continuum) with the less skilled subjects (those at the slow end of the continuum). Individual differences among skilled (faster) subjects at the time of first presentation should reflect a differential use of effortful processes, and therefore CV should correlate positively with RT for the same reasons we saw with the baseline word data from faster subjects. These same subjects should still show gains in perceptual fluency with repeated practice, and so, at the later repetitions, there should still be a positive correlation. This is what we found to be the case: at first presentation, $r = .668$ ($n = 22$, $p < .001$), while at last presentation, $r = .675$ ($n = 22$, $p < .001$).

In contrast, individual differences among less skilled (slower) subjects that were evident at the time of first presentation should not reflect a differential use of effortful processes, since all subjects will be heavily dependent upon them, as was the case with baseline word data for slow subjects. Thus, CV should not correlate strongly with RT for first presentation words. This is exactly what we found: at first presentation, $r = .176$ ($n = 22$, n.s.). Later, however, these same subjects should show gains in perceptual fluency with repeated items, and so, at the time of the last presentation, there should be a positive correlation between CV and RT. This was indeed the case: at last presentation, $r = .507$ ($n = 22, p < .02$).

Thus, in several ways, the results from the second language lexical decision task are consistent with the predictions derived from the theoretical analysis presented earlier. These data show that individual differences and within-subject differences in response latency *SD* can, in an appropriately complex task, reflect changes over and above those expected from simple speed-up of processing. The data are consistent with the idea that, unlike in Experiment 1, faster subjects (presumably those who were more experienced in reading English) initially performed the word recognition task with fewer slow, variable processing components than did the slower subjects. Moreover, with practice, all individuals improved their word recognition performance in a manner consistent with the interpretation that practice improved automaticity.

GENERAL DISCUSSION

This study addressed the questions of whether one can identify differences in automaticity of word recognition as a function of word recognition skill or practice. These questions were addressed directly in the second experiment reported here. Word recognition skill was operationalized in terms of lexical decision latency (with fast responding indicating higher levels of skill). Automaticity was operationalized in terms of a positive correlation between RT and the CV of RT following a theoretical analysis of how a blend of automatic and controlled processes underlying a complex task such as word recognition might contribute to response variability.

The subjects in a second language lexical decision task performed in a manner that confirmed expectations. Between-subject analyses showed (as expected) that CVs correlated positively with RT, indicating that gains in word recognition skill (faster responding) were associated with large decreases in variability of responding. Within-subject analyses also showed (as expected) that all subjects responded faster to items that were repeated within the task, and that CVs for these items correlated positively with RT. Further analyses also revealed that automaticity effects were stronger with skilled (faster) subjects than with less skilled (slower) subjects. Overall, the results conformed nicely to the hypothesized relationships between RT and CV that were derived from a discussion of ways in which controlled and automatic processing help to determine overall response variability.

Such analyses may have interesting and important applications for the study of second language acquisition. For example, it has been argued elsewhere that successful second language training programs should include the goal of automatizing certain fundamental language skills (Gatbonton & Segalowitz, 1988; Segalowitz et al., 1991). Measures of the degree to which a given language skill has been automatized have hitherto been difficult to implement in most situations. The analyses presented in this article show promise in this respect. If one can obtain response latencies in an appropriate context (and one need not be restricted to measures of word recognition) and measures of the variability of those latencies, then it should be possible to investigate whether the CV correlates positively with RT. In this manner, one can study changes in automaticity as a function of different language exposure and training situations.

This study can also be seen to address larger issues related to skill development in domains other than second language acquisition. Consider, for example, what our null hypothesis should be for assessing claims of a gain in automaticity in a given situation. An example might be a task of relatively long duration, such as a music or dance performance of several minutes. This would involve the sequencing of a number of processing units, each with a characteristic execution time and variability. As can be seen from the example in Table 1A, the overall execution time of a hypothetical task involving 10 underlying components appears to be relatively more stable (here, the *SD* of total execution time is 12.7% of the overall execution time) than is the execution time of any one component (here, it is either 20% or 40%, depending on the component). Brothers and Shaw (1989) reported cases in which dancers were able to execute mental performances lasting several minutes with what appeared to be impressively little variability in mean execution time (in the order of a few percentage points). They concluded that there must be some higher order clock that regulates this behavior to permit such precise replicability of performance items. According to the present analysis, however, this may not necessarily be so; it all depends on how many sequential units are presumed to underlie the performance. We have seen that under a null hypothesis that no automatization or higher order mental clocks are involved, one can nevertheless expect there to be a considerable reduction in variability in relation to total performance time. Thus, under some circumstances, it may become necessary to define what counts as an independent, functional unit of processing if one wants to estimate the "normal" reduction in variability before automaticity and higher order clocks enter into the picture.

Finally, it should be noted that this analysis of the difference between speed-up effects and automaticity can be applied to a number of different dependent measures, including latencies of evoked potentials, panel pressing, timing of speech events, eye movements - any situation in which latencies and their variability can be measured over time as a function of practice or skill level. It is hoped that analyses of variability, such as those proposed here, will provide a useful complement to other research techniques for understanding the mechanisms underlying second language performance and complex human skills in general.

ACKNOWLEDGMENTS

This research was supported by a grant from the Quebec Ministry of Education (FCAR). The authors would like to thank Claire Carrière, Elizabeth Gatbonton, Zhenhua Huang, Mel Komoda, Catherine Poulsen, and Vivien Watson for their valuable comments in earlier phases of this research, the Brock University Intensive ESL Program for their cooperation and assistance, and JoAnne Stewart for helping with the data collection. Data from this article were presented at the 25th International Congress of Psychology, Brussels, July 1992.

NOTE

1. The bilinguals in this study were very fluent in their two languages, despite the fact that they read their L2 more slowly than their L1. The means of the 24 bilinguals' RTs were 836 and 707 msec for L2 and L1, respectively, $t(23) = 3.592$, $p < .005$. The means of their individual *SD* of RTs were 310 and 188 msec for L2 and L1, respectively, $t(23) = 3.422$, $p < .005$. The correlation between their RTs and *SD*s in L1 was $r = .742$, $p < .01$ $(n = 24)$. Similarly, for a group of 12 French and 12 English monolinguals in the same study, the mean RT and mean of the individuals' *SD*s were 726 and 222 msec, respectively, $r = .719, p < .01$ $(n = 24)$.

REFERENCES

Ackerman, P. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin, 102*, 3–27.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369–406.

Bekoff, Marc. (1977). Quantitative studies of three areas of classical ethology: Social dominance, behavioral taxonomy and behavioral variability. In B. A. Hazlett (Ed.), *Quantitative methods in the study of animal behavior* (pp. 1–46). New York: Academic.

Brothers, L., & Shaw, G. L. (1989). The role of accurate timing in human performance and the code for higher cortical function. In R. Cotterill (Ed.), *Models of brain function* (pp. 285–291). Cambridge: Cambridge University Press.

Carlton, Mary J., Robertson, R. N., Carlton, Les G., & Newell, K. M. (1985). Response timing variability: Coherence of kinematic and EMG parameters. *Journal of Motor Behavior, 17*, 301-319.

Cheng, P. (1985). Restructuring versus automaticity: Alternative accounts of skill acquisition. *Psychological Review, 92*, 414-423.

Crossman, E. (1959). A theory of the acquisition of speed skill. *Ergonomics, 2*, 153-156.

Dawkins, Richard, & Dawkins, Marian. (1973). Decisions and the uncertainty of behaviour. *Behaviour, 45*, 83-103.

den Heyer, K., Goring, A., Gorgichuk, S., Richards, L., & Landry, M. (1988). Are lexical decisions a good measure of lexical access? Repetition blocking suggest the affirmative. *Canadian Journal of Psychology, 42*, 274-296.

Favreau, Micheline, & Segalowitz, Norman. (1983). Automatic and controlled processes in the first and second language reading of fluent bilinguals. *Memory and Cognition, 11*, 565-574.

Fleishman, E. A., & Mumford, Michael, D. (1989). Abilities as causes of individual differences in skill acquisition. *Human Performance, 2*, 201-223.

Fodor, J. (1983). *Modularity of mind*. Cambridge: MIT Press.

Gatbonton, Elizabeth, & Segalowitz, Norman. (1988). Creative automatization: Principles for promoting fluency within a communicative framework. *TESOL Quarterly, 22*, 473-492.

Guilford, J. P. (1942). *Fundamental statistics in psychology and education* (1st ed.). New York: McGraw Hill.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513-541.

Jacoby, L. L., & Brooks, L. R. (1984). Nonanalytic cognition: Memory, perception, and concept learning. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 18, pp. 1-47). New York: Academic.

Kahneman, Daniel, & Triesman, Anne. (1984). Changing views of attention of automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 29-61). New York: Academic.

Logan, Gordon. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492-527.

(1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology, 22*, 1-35.

MacKay, D. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review, 89*, 483-506.

McManus, I. C., Kemp, R. I., & Grant, J. (1986). Differences between fingers and hands in tapping ability: Dissociation between speed and regularity. *Cortex, 22*, 461-473.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General, 106*, 226-254.

(1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264-336). Hillsdale, NJ: Erlbaum.

Newell, Alan, & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.

Newell, K. M. (1991). Motor skill acquisition. *Annual Review of Psychology, 42*, 213-237.

Perfetti, Charles. (1985). *Reading ability*. New York: Oxford University Press.

Posner, Michael, & Snyder, C. (1975). Attention and cognitive control. In R. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55-85). Hillsdale, NJ: Erlbaum.

Rastatter, Michael, & Blair, Bruce. (1984). Developmental influences on speech-motor equivalence: Some implications for articulatory disordered children. *Perceptual and Motor Skills, 59*, 995-998.

Schneider, W., Dumais, S. T., & Shiffrin, R. (1984). Automatic and control processing and attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 1-27). New York: Academic.

Segalowitz, Norman S. (1986). Skilled reading in the second language. In Jyotsna Vaid (Ed.), *Language processing in bilinguals: Psycholinguistic and neuropsychological perspectives*. Hillsdale, NJ: Erlbaum.

(1991). Does advanced skill in a second language reduce automaticity in the first language? *Language Learning, 41*, 59–83.

Segalowitz, Norman S., & Hébert, Martine. (1990). Phonological recoding in the first and second language reading of skilled bilinguals. *Language Learning, 40*, 59–83.

Segalowitz, Norman S., Poulsen, Catherine, & Komoda, Melvin. (1991). Lower level components of reading skill in higher level bilinguals: Implications for reading instruction. *Revue de l'AILA, 8*, 15–30.

Segalowitz, Sidney J., Unsal, Ayse, & Dywan, Jane. (1992). ERP and behavioral reaction time variability in head injury. *Journal of Clinical and Experimental Neuropsychology, 14*, 73 (abstract).

Shallice, T. (1982). Specific impairments of planning. *Transactions of the Royal Society of London, B298*, 199–209.

Slater, P. J. B. (1978). Data collection. In P. W. Colgan (Ed.), *Quantitative ethology* (pp. 7–24). New York: Wiley.

Sloboda, J. (1985). *The musical mind.* Oxford: Oxford University Press.

Stanovich, Keith E. (1990). Concepts in developmental theories of reading skill: Cognitive resources, automaticity, and modularity. *Developmental Review, 10*, 72–100.

Townsend, J. T. (1974). Issues and models concerning the processing of a finite number of inputs. In Barry Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 133–185). Hillsdale, NJ: Erlbaum.

Watson, John S., & Livesey, Peter J. (1982). Early rearing conditions and within-group behavioral variability. *Behavioral and Neural Biology, 36*, 98–101.

Weismer, Gary, & Elbert, Mary. (1982). Temporal characteristic of "functionally" misarticulated /s/ in 4- to 6-year-old children. *Journal of Speech and Hearing Research, 25*, 275–287.