

Psychological Science

<http://pss.sagepub.com/>

Deliberate Practice Is Necessary but Not Sufficient to Explain Individual Differences in Piano Sight-Reading Skill : The Role of Working Memory Capacity

Elizabeth J. Meinz and David Z. Hambrick

Psychological Science 2010 21: 914 originally published online 9 June 2010

DOI: 10.1177/0956797610373933

The online version of this article can be found at:

<http://pss.sagepub.com/content/21/7/914>

Published by:



<http://www.sagepublications.com>

On behalf of:



[Association for Psychological Science](http://www.sagepublications.com)

Additional services and information for *Psychological Science* can be found at:

Email Alerts: <http://pss.sagepub.com/cgi/alerts>

Subscriptions: <http://pss.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Deliberate Practice Is Necessary but Not Sufficient to Explain Individual Differences in Piano Sight-Reading Skill: The Role of Working Memory Capacity

Elizabeth J. Mainz¹ and David Z. Hambrick²

¹Southern Illinois University Edwardsville and ²Michigan State University

Abstract

Deliberate practice—that is, engagement in activities specifically designed to improve performance in a domain—is strongly predictive of performance in domains such as music and sports. It has even been suggested that deliberate practice is sufficient to account for expert performance. Less clear is whether basic abilities, such as working memory capacity (WMC), add to the prediction of expert performance, above and beyond deliberate practice. In evaluating participants having a wide range of piano-playing skill (novice to expert), we found that deliberate practice accounted for nearly half of the total variance in piano sight-reading performance. However, there was an incremental positive effect of WMC, and there was no evidence that deliberate practice reduced this effect. Evidence indicates that WMC is highly general, stable, and heritable, and thus our results call into question the view that expert performance is solely a reflection of deliberate practice.

Keywords

expertise, individual differences, working memory, music

Received 11/8/08; Revision accepted 12/1/09

For more than a century, scientists have debated the origin of expertise. Galton (1869) argued that genius arises from inherited ability, but more recent accounts tend to emphasize the role of experience. Most prominently, Ericsson and Ward (2007) documented that expertise is attained only after 10 years of intense preparation, and Ericsson and Charness (1994) argued that “basic abilities and capacities (talent) that remain stable in studies of limited short-term practice do not generalize to superior performance acquired over years and decades in a specific domain” (p. 731). According to this view, expert performance is mediated by acquired mechanisms, which are assumed to “relieve” basic capacities as predictors of performance. As Roring, Nandagopal, and Ericsson (2007) stated, “during years of practice and training, experts acquire elaborate mechanisms for encoding and maintaining flexible access to critical task information that bypass basic capacities, such as short-term memory capacity” (p. 169).

This view is supported by the finding that estimates of deliberate practice, which is defined as cumulative engagement in activities specifically designed to improve performance in a domain, correlate positively with measures of skill in domains such as chess (Charness, Tuffiash, Krampe, Reinhold, & Vasyukova, 2005) and music (Ruthsatz, Detterman,

Griscom, & Cirullo, 2008), and in everyday activities such as typing (Keith & Ericsson, 2007). This evidence indicates that deliberate practice is necessary for acquiring expertise. But if two individuals practice an equal amount with equal intensity, will they reach the same skill level?

Ericsson and his colleagues reported nonsignificant relationships between measures of cognitive ability and skill in piano playing (Krampe & Ericsson, 1996), typing (Keith & Ericsson, 2007), and the word game Scrabble (Tuffiash, Roring, & Ericsson, 2007), although Grabner, Stern, and Neubauer (2007) found that IQ predicted chess skill even after taking practice into account. Ruthsatz et al. (2007) found that Raven’s Progressive Matrices scores correlated positively and significantly with musical achievement in high-school band members, but not in conservatory students. However, Ruthsatz et al. did not test whether this difference in correlations was due to increasing deliberate practice (i.e., the Deliberate Practice \times Ability interaction). Overall, there is little evidence

Corresponding Author:

Elizabeth J. Mainz, Department of Psychology, Southern Illinois University Edwardsville, Box 1121, Edwardsville, IL 62026-1121
E-mail: emeinz@siue.edu

directly relevant to the question of whether deliberate practice eliminates the effects of basic abilities on domain-relevant performance.

Working Memory Capacity

Working memory capacity (WMC) is the ability to maintain task-relevant information in a highly active state. It is typically measured with complex span tasks, such as operation span (Engle, 2002), in which the participant solves equations while remembering words. Scores on such tasks correlate with success in complex cognitive tasks and with each other (Kane et al., 2004), which suggests that WMC is highly general. Furthermore, WMC is influenced substantially by genetic factors, with heritability estimates around 50% (e.g., Kremen et al., 2007; Polderman et al., 2006), and is stable across time (e.g., Unsworth, Heitz, Schrock, & Engle, 2005).

Few studies have investigated the role of WMC in expert performance, but there are suggestions it plays an important role. Robbins et al. (1996) found that a secondary task (random-letter generation) designed to tax WMC impaired performance in a chess-move task, even in highly skilled players, and Hambrick and Engle (2002) found that WMC predicted participants' ability to keep track of events in fictitious baseball games, even among participants with a high level of baseball knowledge.

The Effect of WMC on Sight-Reading

Pianists are sometimes called on to sight-read—to play music with little or no preparation, for example, when accompanying a soloist. Just as typists look ahead in text to prepare keystrokes (Salthouse, 1984), skilled pianists look ahead in music scores during sight-reading (Rayner & Pollatsek, 1997). Indeed, Sloboda (1974) reported a near-perfect correlation between sight-reading skill and the eye-hand span (i.e., the number of notes pianists could play in a piece after it was made invisible). WMC may affect sight-reading performance; for example, a high-WMC pianist should be able to maintain more notes beyond the ones currently being played than a low-WMC pianist can.

In a study by Kopiez and Lee (2006, 2008), pianists across a wide range of sight-reading skill performed a sight-reading task with five difficulty levels. For all but one level, a measure of WMC correlated positively and significantly with performance (collapsing across levels, the correlation was just shy of significance, $r = .26$). Thus, there is some evidence suggesting that WMC contributes to individual differences in sight-reading skill. Our question was whether this is true even at high levels of deliberate practice.

In our study, pianists across a wide range of skill completed a test of sight-reading, as well as tests and questionnaires assessing deliberate practice and WMC. We then performed regression analyses to test for main and interactive effects of deliberate practice and WMC on sight-reading performance.

Method

Participants

Participants ($N = 57$) were recruited through newspaper advertisements, professional contacts, and a university music department. They were paid \$40, except for 4 undergraduates, who received course credit. The sample was predominantly female (73.7%) and under 50 years old ($M = 30.9$ years). Nearly half (47.4%) of participants held one or more music degrees; 80.7% considered piano their primary instrument. Other characteristics of the participants are summarized in Table 1. Overall piano-playing experience ranged from 1 to 57 years; cumulative deliberate practice ranged from 260 to more than 31,000 hr (median = 4,160).

Materials and procedure

Participants were tested individually in 3-hr sessions; materials were administered in a fixed order.

Experience questionnaire and interview. A background questionnaire included questions about primary and secondary musical instruments, professional experiences, and lessons. To gather estimates of deliberate practice (performed alone), we interviewed participants about their piano-playing history. Following the procedures used by Ericsson, Krampe, and Tesch-Römer (1993), we had participants indicate on a time line when lessons began and ended, when teachers changed, and other events, to provide reference points for estimates. Participants then estimated, for each year, the number of hours per week they spent on deliberate practice alone, with the goal of improving their performance (not simply playing for enjoyment); participants also estimated the number of hours they spent on sight-reading practice. Weekly estimates of deliberate practice were multiplied by 52 and summed to create measures of cumulative deliberate practice.

Sight-reading. Before performing, participants familiarized themselves with the electronic piano we used in the study: a full-size keyboard (Yamaha Portable Grand DGX-500) with 88 weighted keys and a damper pedal. The sight-reading test consisted of six pieces, all from the *Four Star Sight Reading and Ear Tests* series (Berlin & Markow, 2002). A piano-pedagogy expert (H.L. Chin, personal communication, May, 2003) recommended this series because it includes pieces of various difficulty levels and was not likely to be familiar to participants, as it is not widely used for evaluation in the United States. We selected excerpts from pieces in Book 6 (low difficulty), Book 8 (medium difficulty), and Book 10 (high difficulty). For each level, one piece was chosen for its level of technical difficulty, and another was chosen for the degree to which it allowed artistic expression. For each piece, recommended tempo was indicated using a standard descriptive term (e.g., Allegretto) and a metronome marking (e.g., ♩ = 116).

Table 1. Participants' Characteristics and Correlations With Sight-Reading Performance

Variable	<i>M</i>	Median	<i>SD</i>	Range	Correlation with sight-reading performance (<i>r</i>)
Years of piano-playing experience	18.6	15	14.3	1–57	.37**
Paid engagements ^a	97.8	4	267.6	0–1,500	.40**
Memorized solos ^a	25.2	5	105.5	0–800	.33*
Accompaniments ^a	70.1	4	280.1	0–2,000	.63**
Years of lessons	10.0	10	5.3	2–23	.61**
Hours of deliberate sight-reading practice ^a	1,486.9	1,040	1,517.0	0–9,048	.48**
Hours of overall deliberate practice ^a	5,805.9	4,160	5,946.8	260–31,096	.67**
Working memory capacity	17.4	17.5	5.3	6.8–33.5	.28*
Sight-reading performance	4.7	4.8	1.9	0.8–7.0	—

Note: *N* = 57. Working memory capacity was calculated as the average of scores for operation span, reading span, rotation span, and matrix span. Sight-reading performance was calculated as the average of the overall ratings across raters, attempts, and difficulty levels.

^aThis variable was nonnormal (skewness > 2) and was log-transformed (log 10), by adding a small constant to remove zeros, before computing its correlation with sight-reading performance.

p* < .05. *p* < .01.

Participants were given 60 s to review a piece. They were then asked to play the piece without stopping or repeating incorrectly performed passages. After their first attempt, participants were given another 60 s to review the piece, after which they played it again. Participants performed the low-, medium-, and then high-difficulty pieces. This format is typical of auditions for university music programs and for evaluations of students in private studios (H.L. Chin, personal communication, April, 2008). Performances were audio-recorded for evaluation. Using a scale from 1 (lowest) to 7 (highest), two expert raters, each of whom taught piano at the university level and held a graduate degree in music, independently and blindly rated each performance on three dimensions: technical proficiency, musicality, and overall performance. (A score of 0 was assigned in a few cases when a piece was too difficult for a particular pianist to attempt.)

WMC. Four tasks, each of which had a decision component and a memory component, were used to measure WMC; all of these tasks have been used previously (e.g., Hambrick & Oswald, 2005). In the **operation span task**, each trial included an equation with an answer, and a word. The participant was instructed to say “yes” if the answer was correct or “no” if the answer was incorrect, and to remember the word. In the **reading span task**, the equations were replaced with sentences requiring participants to respond “yes” or “no” to indicate whether the sentences made sense. In the **rotation span task**, each trial consisted of a rotated letter and an arrow pointing away from the letter. The participant was instructed to say “yes” if the rotated letter was oriented normally or “no” if it was displayed as a mirror image, and to remember the direction of the arrow. In the **matrix span task**, the stimuli for the processing component were 4 × 4 arrays of the letter *L*; these arrays could include a target (⊥). The participant was instructed to say “yes” if the target was present or “no” if it was not. A

different matrix that had one blue cell then appeared, and the participant was instructed to remember that cell’s location. In each task, after responding to three to six pairs of components, the participant was instructed to record the items to be remembered on a response sheet in the order in which they had been presented; 1 point was awarded for both correct recall in the memory component and a correct judgment in the yes/no decision component, and the score was the total number of points.

Power analyses. Power analyses indicated that 50 participants would provide adequate power (> .80) in a regression analysis to detect a medium-sized interaction effect (Cohen’s $f^2 = .15$) between two predictor variables.

Results

Ten participants completed only two WMC tasks, so some WMC values, along with a small amount of other data, were missing (< 1% of the data). We replaced all missing values using the Expectation-Maximization algorithm (SPSS Version 15.0); we also screened for values more than 3.5 standard deviations from sample means (outliers) and replaced the two values that met this criterion with the cutoff value. All variables used in subsequent analyses were approximately normally distributed, except for both measures of cumulative deliberate practice (skewness > 2). As other researchers have done (e.g., Charness et al., 2005), we log-transformed these variables.

The four WMC variables correlated moderately with each other (r s = .35–.77), as expected. For sight-reading performance, correlations between raters were acceptably high (average $r = .86$). The overall, technical, and musical ratings correlated highly ($r > .70$), as did the overall ratings across attempts and difficulty levels ($r > .75$). Thus, we created a WMC composite by averaging WMC scores and a measure of sight-reading by averaging overall ratings across raters,

attempts, and difficulty levels. Coefficient alphas were .81 for WMC and .99 for sight-reading performance, indicating good reliability.¹

As shown in Table 1, there was a strong positive correlation between overall deliberate practice and sight-reading performance ($r = .67, p < .01$); this correlation was essentially unchanged after controlling for WMC (partial $r = .70, p < .01$). However, there was also a positive correlation between WMC and sight-reading performance ($r = .28, p < .05$); this correlation remained significant, and even increased slightly, after controlling for deliberate practice (partial $r = .37, p < .01$). (The correlation between WMC and deliberate practice was near zero.)

To answer our primary question, we performed a hierarchical regression analysis predicting sight-reading performance. Overall deliberate practice was entered in Step 1, WMC in Step 2, and the Deliberate Practice \times WMC interaction in Step 3. Statistically significant increments in variance (ΔR^2 's) would indicate contributions of WMC to sight-reading performance, above and beyond deliberate practice (Step 1 to Step 2), and a Deliberate Practice \times WMC interaction (Step 2 to Step 3).

Results of this analysis are shown in Table 2. Deliberate practice accounted for a large proportion of the variance in sight-reading performance, $R^2 = .451, F(1, 55) = 45.24, p < .01$. Clearly, high levels of deliberate practice were associated with superior sight-reading performance. However, there was a moderate effect of WMC (Cohen, 1988), above and beyond deliberate practice, $R^2 = .074, F(1, 54) = 8.48, p < .01$. Furthermore, the Deliberate Practice \times WMC interaction was nonsignificant ($F < 1$). High levels of WMC were associated with superior sight-reading performance regardless of the level of deliberate practice.

We repeated the analysis, entering age in Step 1 along with deliberate practice. The incremental effect of WMC on sight-reading performance was nearly identical to that in our first analysis ($R^2 = .074, p < .01$). In a third analysis, we entered a perceptual-speed composite in Step 2 along with WMC.² The unique effect of WMC on sight-reading performance ($R^2 = .053, p < .05$) was significant, whereas the effect of speed on sight-reading performance was nil ($F < 1$). The effect of WMC on sight-reading performance was not influenced by age or by the demands of speeded responding in the sight-reading task.

To investigate whether deliberate practice devoted to sight-reading would predict performance better than broader indices of deliberate practice (as suggested by Kopiez & Lee, 2008;

Lehmann & Ericsson, 1996), we performed another regression analysis. Sight-reading practice accounted for a significant proportion of variance in sight-reading performance, $R^2 = .228, F(1, 55) = 16.21, p < .01$, but R^2 increased by .224 to .451, $F(1, 54) = 22.03, p < .01$, after overall deliberate practice was entered into the model. This increase suggests that sight-reading skill arises not only from practice devoted specifically to this activity, but also from other practice activities. Finally, above and beyond the effect of sight-reading practice, the effect of WMC was still significant, $R^2 = .075, F(1, 53) = 8.34, p < .01$, and there was no Sight-Reading Practice \times WMC interaction ($F < 1$). High levels of WMC were associated with superior sight-reading performance regardless of level of sight-reading practice.

Discussion

Across a wide range of piano-playing skill, deliberate practice accounted for nearly half the variance (45.1%) in sight-reading performance. However, WMC accounted for a significant proportion of the variance (7.4%), above and beyond deliberate practice, and there was no evidence that deliberate practice reduced this effect. We cannot rule out the possibility that there is some degree of practice-related reduction of the effect of WMC on sight-reading performance. However, our results suggest that this reduction is likely to be small: A sample size of more than 20,000 subjects would be required to detect a Deliberate Practice \times WMC interaction of the magnitude we observed ($R^2 < .0001$) as statistically significant with acceptable power.

Schellenberg (2004) reported that music training enhanced IQ in children who had low levels of music experience, and there is one report that music training may enhance WMC (Lee, Lu, & Ko, 2007). Nevertheless, there is no indication that our results reflect enhancement of WMC through piano training: WMC had a near-zero correlation with both overall deliberate practice and sight-reading practice. We speculate that WMC plays a direct role in sight-reading performance by determining the extent to which pianists can prepare for future keystrokes by looking ahead in music scores.

WMC is highly general, stable, and heritable. Our results therefore call into question the view, advocated by Ericsson and his colleagues, as well as other researchers (e.g., Howe, Davidson, & Sloboda, 1998), that basic capabilities such as

Table 2. Results of the Hierarchical Regression Analysis Predicting Sight-Reading Performance

Step	ΔR^2	ΔF	df	β	p_{rep}
1. Deliberate practice	.451 (.441)	45.24**	1, 55	0.672	> .99
2. Working memory capacity	.074 (.067)	8.48**	1, 54	0.273	.97
3. Deliberate Practice \times Working Memory Capacity	.000 (.000)	0.00	1, 53	-0.005	.11

Note: $N = 57$. Values in parentheses are "shrunken" R^2 's, which are adjusted for the number of predictor variables and sample size.

** $p < .01$.

WMC are unimportant for expert performance. Although it seems reasonable to predict that anyone who engages in thousands of hours of deliberate practice will develop a high level of sight-reading skill, WMC may limit the ultimate level of performance that can be attained. More generally, we suggest that deliberate practice—although necessary for acquiring expertise—will not always be sufficient to overcome limitations due to basic abilities.

Acknowledgments

The authors made equal contributions to this project. We thank Vera Parkin, Hwei Li Chin, Sherri Downey, and Emily Severine. We also thank Kim Fenn for comments.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This project was funded by a Funded University Research grant to E. Meinz from Southern Illinois University Edwardsville.

Notes

1. We computed reliability and variability of the sight-reading variable separately for the lower, middle, and upper terciles of deliberate practice. There was no evidence for differential sensitivity across the range of deliberate practice, as coefficient alphas were the same for these groups (all α s < .93), and standard deviations were similar among the groups (lower = 1.8, middle = 1.3, upper = 1.3). There is also no indication of a ceiling effect in sight-reading performance, as it was approximately normally distributed (skewness < -1).
2. We measured perceptual speed with four tasks requiring speeded judgments of letter pairs, number pairs, pictures, and patterns (see Hambrick & Oswald, 2005).

References

- Berlin, B., & Markow, A. (2002). *Four star sight reading and ear tests: Daily exercises for piano students* (Books 6, 8, and 10; S. McBride Smith, Series Ed.). Mississauga, Ontario, Canada: Frederick Harris.
- Charness, N., Tuffiash, M., Krampe, R.T., Reingold, E., & Vasyukova, E. (2005). The role of deliberate practice in chess expertise. *Applied Cognitive Psychology*, 19, 151–165.
- Cohen, J. (1988). *Statistical power for the behavioral sciences* (2nd ed.). Mahwah, NJ: Erlbaum.
- Engle, R.W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, 11, 19–23.
- Ericsson, K.A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, 49, 725–747.
- Ericsson, K.A., Krampe, R.T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100, 363–406.
- Ericsson, K.A., & Ward, P. (2007). Capturing the naturally occurring superior performance of experts in the laboratory. *Current Directions in Psychological Science*, 16, 346–350.
- Galton, F. (1869). *Hereditary genius*. London: Macmillan.
- Grabner, R.H., Stern, E., & Neubauer, A.C. (2007). Individual differences in chess expertise: A psychometric investigation. *Acta Psychologica*, 124, 398–420.
- Hambrick, D.Z., & Engle, R.W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. *Cognitive Psychology*, 44, 339–387.
- Hambrick, D.Z., & Oswald, F.L. (2005). Does domain knowledge moderate involvement of working memory capacity in higher-level cognition? A test of three models. *Journal of Memory and Language*, 52, 377–397.
- Howe, M.J.A., Davidson, J.W., & Sloboda, J.A. (1998). Innate talents: Reality or myth? *Behavioral & Brain Sciences*, 21, 399–442.
- Kane, M.J., Hambrick, D.Z., Tuholski, S.W., Wilhelm, O., Payne, T.W., & Engle, R.W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, 133, 189–217.
- Keith, N., & Ericsson, K.A. (2007). A deliberate practice account of typing proficiency in everyday typists. *Journal of Experimental Psychology: Applied*, 13, 135–145.
- Kopiez, R., & Lee, J.I. (2006). Towards a dynamic model of skills involved in sight reading music. *Music Education Research*, 8, 97–120.
- Kopiez, R., & Lee, J.I. (2008). Towards a general model of skills involved in sight reading music. *Music Education Research*, 10, 41–62.
- Krampe, R.T., & Ericsson, K.A. (1996). Maintaining excellence: Deliberate practice and elite performance in younger and older pianists. *Journal of Experimental Psychology: General*, 125, 331–359.
- Kremen, W.S., Jacobsen, K.C., Xian, H., Eisen, S.A., Eaves, L.J., Tsuang, M.T., & Lyons, M.J. (2007). Genetics of verbal working memory processes: A twin study of middle-aged men. *Neuropsychology*, 21, 569–580.
- Lee, Y., Lu, M., & Ko, H. (2007). Effects of skill training on working memory capacity. *Learning and Instruction*, 17, 336–344.
- Lehmann, A.C., & Ericsson, K.A. (1996). Performance without preparation: Structure and acquisition of expert sight-reading and accompanying performance. *Psychomusicology*, 15, 1–29.
- Polderman, T.J.C., Stins, J.F., Posthuma, D., Gonso, M.F., Verhulst, F.C., & Boomsma, D.I. (2006). The phenotypic and genotypic relation between working memory speed and capacity. *Intelligence*, 34, 549–560.
- Rayner, K., & Pollatsek, A. (1997). Eye movements, the eye-hand span, and the perceptual span during sight-reading of music. *Current Directions in Psychological Science*, 6, 49–53.
- Robbins, T.W., Henderson, E.J., Barker, D.R., Bradley, A.C., Fearneough, C., Henson, R., et al. (1996). Working memory in chess. *Memory & Cognition*, 24, 83–93.
- Roring, R.W., Nandagopal, K., & Ericsson, K.A. (2007). Can Parieto-Frontal Integration Theory be extended to account for individual differences in skilled and expert performance in everyday life? *Behavioral & Brain Sciences*, 30, 168–169.
- Ruthsatz, J., Detterman, D., Griscom, W.S., & Cirullo, B.A. (2008). Becoming an expert in the musical domain: It takes more than just practice. *Intelligence*, 36, 330–338.

- Salthouse, T.A. (1984). Effects of age and skill in typing. *Journal of Experimental Psychology: General*, 13, 345–371.
- Schellenberg, E.G. (2004). Music lessons enhance IQ. *Psychological Science*, 15, 511–514.
- Sloboda, J.A. (1974). The eye-hand span: An approach to the study of sight reading. *Psychology of Music*, 2, 4–10.
- Tuffiash, M., Roring, R.W., & Ericsson, K.A. (2007). Expert word play: Capturing and explaining reproducibly superior verbal task performance. *Journal of Experimental Psychology: Applied*, 13, 124–134.
- Unsworth, N., Heitz, R.P., Schrock, J.C., & Engle, R.W. (2005). An automated version of the operation span task. *Behavior Research Methods*, 37, 498–505.