

Training specificity and transfer in time and distance estimation

Alice F. Healy · Lindsay Anderson Tack ·
Vivian I. Schneider · Immanuel Barshi

Published online: 24 January 2015
© Psychonomic Society, Inc. 2015

Abstract Learning is often specific to the conditions of training, making it important to identify which aspects of the testing environment are crucial to be matched in the training environment. In the present study, we examined training specificity in time and distance estimation tasks that differed only in the focus of processing (FOP). External spatial cues were provided for the distance estimation task and for the time estimation task in one condition, but not in another. The presence of a concurrent alphabet secondary task was manipulated during training and testing in all estimation conditions in Experiment 1. For distance as well as for time estimation in both conditions, training of the primary estimation task was found to be specific to the presence of the secondary task. In Experiments 2 and 3, we examined transfer between one estimation task and another, with no secondary task in either case. When all conditions were equal aside from the FOP instructions, including the presence of external spatial cues, Experiment 2 showed “transfer” between tasks, suggesting that training might not be specific to the FOP. When the external spatial cues were removed from the time estimation task, Experiment 3 showed no transfer between time and distance estimations, suggesting that external task cues influenced the procedures used in the estimation tasks.

Keywords Skill acquisition · Time estimation · Distance estimation · Training specificity · Transfer

A. F. Healy (✉) · L. A. Tack · V. I. Schneider
Department of Psychology and Neuroscience, University of
Colorado, Muenzinger Building, 345 UCB,
Boulder, CO 80309-0345, USA
e-mail: alice.healy@colorado.edu

I. Barshi
Human Systems Integration Division, NASA, Ames Research
Center, Moffett Field, Mountain View, CA, USA

How far from your home is the nearest airport? There are two ways to answer this question: one in terms of the time it takes to get there (e.g., 45 min), and the other in terms of the distance between the locations (e.g., 45 miles). These two perspectives can be viewed as different in the focus of processing (FOP). In the present study, we explore differences in the training of time and distance estimates, to see how specific is the learning in each case. Can training of time estimates transfer to making distance estimates, and vice versa, or is training specific to the FOP? Also, are both types of estimates dependent on the context in which they are made; that is, do they both depend on the presence of a secondary task?

Learning is often highly specific to the conditions under which training took place, limiting the transfer of trained knowledge to new situations, environments, or novel instances. Such specificity has been found in a variety of procedural skills, including the Stroop color-word interference task, in which subjects responded faster at test when presented with a trained color-word set than with a new color-word set (Clawson, King, Healy, & Ericsson, 1995); tasks involving mental arithmetic, in which subjects responded faster at test to trained problems than to the same problems with a change in operation or in operand order (Rickard, Healy, & Bourne, 1994); and tasks requiring speeded aiming movements with defective computer mice, in which subjects who learned how to cope with one type of defective mouse showed interference when they encountered a different type of defective mouse (Healy, Wohldmann, Sutton, & Bourne, 2006), or even when they returned to a mouse with no defects (Healy, Wohldmann, & Bourne, 2011). These investigators all explained the observed specificity of training in terms of the procedural reinstatement principle (e.g., Lohse & Healy, 2012), according to which test performance is best when the mental procedures developed in training can be used in subsequent testing. This principle was derived from the procedural view espoused by Kolers and Roediger (1984), from demonstrations of transfer-

appropriate processing (McDaniel, Friedman, & Bourne, 1978; Morris, Bransford, & Franks, 1977; Roediger, Weldon, & Challis, 1989) and encoding specificity (Tulving & Thomson, 1973), and from identical-elements models of learning (e.g., Rickard & Bourne, 1996; Singley & Anderson, 1989; Thorndike, 1906).

In contrast to the high specificity observed in these studies documenting the procedural reinstatement principle are investigations that have shown reliable transfer of skills from one task to another. For example, in a recent meta-analysis of training studies involving spatial skills (Uttal et al., 2013), it was concluded that training of a given spatial task transferred to other spatial tasks that were not directly trained. In fact, most impressive along these lines have been studies showing that working memory training can lead to improvements on tests of fluid intelligence, even though the trained tasks are very different from the intelligence tests (see, e.g., Jaeggi, Buschkuhl, Jonides, & Perrig, 2008, and Sternberg, 2008; but see also Shipstead, Redick, & Engle, 2012, for a cautionary review of these claims).

Because of the previous findings involving the transfer of training for spatial skills, in the present study we sought to determine whether the specificity of training found earlier for various other procedural skills would also apply to a task involving spatial skills. Although the degree of transfer and of specificity of skills undoubtedly depends on the exact nature of what is being learned in a given task, according to the procedural reinstatement principle, there is no reason why spatial skills should show any less specificity of training than other types of procedural skills.

Perhaps the strongest evidence for specificity of training and lack of transfer of training has come from previous research using a dual-task methodology (Healy, Wohldmann, Parker, & Bourne, 2005), in which subjects completed two tasks concurrently, one as the primary task and the other as a secondary task. It was found that primary-task training led to improved test performance on the primary task only when the secondary-task conditions at test matched those at training. In that study, no transfer of training of the primary task occurred when there was a change from training to testing in the secondary-task conditions, even though the training and testing both involved that primary task. The secondary task in that study required subjects to recite aloud the alphabet backward by every third letter, in one case, or to do nothing, in the other case. The primary task used in that study involved time estimation; specifically, subjects produced time intervals of a specified length by pressing a response key when the elapsed time after a beep was equal to the target time. It is possible that the strong specificity found in that case was an artifact of the time estimation task, because the secondary task filled the time interval being estimated. For example, subjects seemed to integrate the primary time estimation and secondary alphabet recitation tasks into a single functional task by using the

number of letters recited as a gauge of how much time had elapsed (see Wohldmann, Healy, & Bourne, 2010, 2012).

To see whether the reliance of the primary-task performance on the secondary task depends on the nature of the primary task, in Experiment 1 we replicated the earlier study with the addition of a new task involving distance estimation, and we compared the two different but related estimation tasks—time and distance estimation. In both of these tasks, the subjects' goal is essentially to learn arbitrary units (temporal or spatial) devised by the experimenter, and these units are the same for the time and distance estimation tasks. The time estimation task is equivalent to the one used earlier, whereas the distance estimation task requires subjects to produce line lengths of a given magnitude by indicating when a stimulus line growing at a fixed rate has reached a certain length. This distance estimation task was chosen in part because it involves spatial skills, which were found to exhibit high transfer in the meta-analysis of Uttal et al. (2013). The distance estimation task was also selected because it could be closely aligned with the time estimation task in almost all respects, except for the task focus on spatial, rather than temporal, information. Substantial differences have been observed in the types of coding used for temporal and spatial information in short-term memory (e.g., Healy, 1975), and such differences might well be associated with differences in the learning of temporal and spatial skills.

In each estimation task group of Experiment 1, half of the subjects were trained to perform the primary estimation task concurrently with the secondary alphabet task, and the remaining subjects were trained to perform the primary estimation task alone in silence. All subjects were tested with the same primary estimation task used at training, but half of the subjects in each estimation task group had the same secondary-task condition as at training, whereas the secondary-task condition was changed for the other half of the subjects in each estimation task group. If the strong specificity observed in the earlier study was an artifact of the time estimation task requirements, such specificity of training should be found only with that task. In contrast, if training is more generally specific to the secondary-task requirements, as implied by the procedural reinstatement principle, then the same pattern of results should be found for distance estimation as for time estimation. For example, subjects could integrate the primary estimation and secondary alphabet recitation tasks to form a single functional task that uses information about how many letters have been recited in the secondary task to gauge not only how much time has elapsed, but also how long the stimulus line has grown.

Experiment 1 included one distance estimation task and two time estimation tasks. The first time estimation task and the distance estimation task differed only in their FOPs, on either time or distance. Specifically, subjects were shown a stimulus line that continually grew in length at a fixed rate,

with the start of the line varying in terms of its horizontal location. For this time estimation task, the subjects were told to press a key after a certain time interval had passed, and for the distance estimation task, the subjects were told instead to press a key after the stimulus line had reached a certain length. These two estimation tasks were equivalent, such that the required ending time equaled the required ending distance (e.g., 20 units of distance is the same as 20 units of time, in terms of the distance that the line grows in that amount of time). Finding less training specificity to the secondary task for the distance estimation task than for this time estimation task under these controlled conditions would allow us to conclude that the magnitude of training specificity depends on the FOP. The second time estimation task examined in Experiment 1 was similar to that used by Healy et al. (2005). In that task, time estimation occurred without the presence of any growing line and was otherwise identical to the other time estimation condition. Including this condition allowed for assessing whether any observed differences between time and distance estimation in their specificity to the secondary task were due to the presence of an external stimulus line.

In Experiment 1, each subject performed only one of the three estimation tasks, and we examined transfer across different secondary-task conditions. In contrast, there was no variation in secondary-task conditions in Experiments 2 and 3; instead, in these experiments we examined transfer across different primary-task conditions—that is, between one primary estimation task and another, subsequent primary estimation task. Does training with time estimation enhance performance on distance estimation, and vice versa, or is training specific to the FOP? Uttal et al. (2013) found transfer from one spatial task to another, but in the present experiments we considered whether transfer takes place from a spatial task to a similar temporal task, or vice versa. To address this issue, we varied both the training condition (time estimation training, distance estimation training) and the testing condition (time estimation test, distance estimation test) in Experiments 2 and 3. Even though the conditions of training on the two tasks should be comparable, the differences in FOPs might preclude transfer across tasks, so that testing on a given task would not benefit from prior training on the alternate task, relative to no prior training. Such a finding would indicate that training specificity applies even to the FOP. In Experiment 2, the time estimation task, like the distance estimation task, included the presence of a growing line, whereas no growing line was present in the time estimation task of Experiment 3.

Experiment 1

The subjects were assigned to one of three training and testing conditions, depending on what they estimated and what they

saw. In the *distance* estimation condition, they estimated the distance of a line that they saw continuously growing across the screen, with their task being to indicate when the line had reached the target length in terms of the number of distance units. In the time estimation conditions, subjects instead estimated time. In one condition (*time line*), they also saw the continuously growing line on the screen, and in the other condition (*time*), no line was shown on the screen. In both of those conditions, the subjects' task was to indicate when a given number of time units had passed. Within each of these three estimation conditions, subjects were in one of four combinations of training and testing secondary-task conditions, depending on whether or not they conducted a difficult alphabet secondary task concurrently with the time estimation task. This design allowed for the determination in each of the three estimation conditions of whether maintaining the same secondary-task condition at training and testing was advantageous for estimation accuracy at test, relative to switching the secondary task between training and testing, which would reflect the amount of specificity of training for the primary estimation task that was dependent on the presence of the secondary task.

Method

Subjects A total of 156 undergraduates from an introductory psychology course at the University of Colorado participated for course credit. There were 12 subject groups, with 13 subjects in each group. Subjects were assigned to groups by a fixed rotation based on their time of arrival for testing.

Design The 12 subject groups varied as a function of estimation condition (distance, time, or time line), secondary-task continuity (same or switch), and secondary-task presence at test (no or yes). A mixed factorial design was, thus, used with three between-subjects variables—estimation condition, secondary-task continuity, and secondary-task presence at test—and two within-subjects variables—phase (training or testing) and block of trials (1–6). The dependent variable was the proportional absolute error, which is the absolute difference between the interval produced and the specified interval divided by the specified interval [i.e., $(|\text{produced interval} - \text{specified interval}|) / \text{specified interval}$], which provides a normalized assessment of the magnitude of the estimation error.

Procedure Subjects were told to estimate a given number of units of either a line distance interval or a time interval. For the time interval, they were told that they would “be asked to estimate a certain interval of time, expressed in fixed arbitrary ‘units’ (not standard units like seconds or minutes).” For the distance interval, they were told that they would “be asked to estimate a certain interval of line distance, expressed in fixed arbitrary ‘units’ (not standard units like centimeters or

inches).” Specifically, the subjects’ task was to produce a specified number of units (time or distance) for six blocks of six trials. On a given trial, subjects were shown the number of units for them to estimate (e.g., “After the beep, estimate x units.”). Next, they heard a beep and saw a cross or a random letter cue in a random location somewhere along the middle row of the screen.

For the distance and time line conditions, subjects saw a growing line that began at the cross or letter cue, increased toward the right, and continued increasing in length at a fixed rate until it reached the right edge of the screen, after which it continued one row below, from the left edge of the screen increasing toward the right. No mention of this growing line was made in the instructions for the time line condition. When the subjects thought that the target distance or time had been attained, they were to press the space bar to indicate that the specified number of units had been reached since the beep had been heard and the cross or letter had appeared on the screen. After making their space-bar response, subjects received immediate feedback, which stated the subjects’ estimated interval length and the difference between their estimated interval and the specified target interval (e.g., “Your estimate was x units. This was y units too short.”). Feedback was provided on every trial of training and testing.

Each block of training and testing had one trial of six different target intervals (21, 25, 32, 47, 50, and 56). For each subject, the order of target intervals was random in a given phase but remained the same for all six blocks in that phase. Unknown to the subject was the fact that each time unit was equal to 783 ms (as in the study by Healy et al., 2005). The unit of line length corresponded to that of the time unit, so that correct responses would be the same in the two conditions. Subjects learned the information about unit length indirectly throughout training and testing via the feedback that they received after making each response. The training and testing phases were separated by a timed unfilled break of 5 min.

The secondary task, which was the same as the difficult alphabet task used by Healy et al. (2005), consisted of reciting aloud the alphabet backward by every third letter (i.e., skipping two letters), starting with the given random letter cue (e.g., if the cue was “t,” the subjects would say “t, q, n, k,” etc.). Subjects were told that if they reached the start of the alphabet (a) in their recitation, they were to revert to the end of the alphabet (z) and continue from that point. They performed this secondary task during the entire interval from the presentation of the letter cue to their response of pressing the space bar. No feedback was given to subjects concerning the accuracy of these secondary-task responses, and although they were not scored, the experimenter was in the room with each subject to monitor the subject’s responses. Also, if the subjects made a procedural error (e.g., recited forward instead of backward or did not skip any letters), the experimenter gave the subjects general feedback about their incorrect procedure. To

ensure that the subjects could perform the secondary alphabet task correctly, before the first phase containing the secondary task the subject was instructed to practice for the experimenter the task of reciting aloud the alphabet in backward order by every third letter, beginning from the letter “m,” and was prompted to continue to do so until he or she reached the letter “x” (i.e., m, j, g, d, a, x). Specific feedback was given to the subject during this initial practice of the secondary task.

Results

Overall analysis An initial overall mixed factorial analysis of variance was conducted to examine the acquisition and retention of the estimation skill across the six blocks of training and testing. The analysis included the between-subjects factors Estimation Condition, Secondary-Task Presence at Test, and Secondary-Task Continuity, along with the within-subjects factors Phase and Block (see Table 1 for a statistical summary of the results). Here and in subsequent analyses, the measure of performance was proportional absolute error. All significant effects ($p < .05$) are reported in the text.

Of greatest interest was the effect of secondary-task continuity (same, switch), which should be evident in the testing phase but not in the training phase if training is specific to the secondary task. However, in previous studies using the time estimation task (e.g., Healy et al., 2005; Wohldmann et al., 2010, 2012), most of the learning occurred in the first block of training trials, so the crucial predicted interaction of secondary-task continuity and phase might only be evident when considering the first block of trials (i.e., a significant three-way interaction of secondary-task continuity, phase, and block might occur). Indeed, as in the earlier studies with this task, we observed a significant interaction of phase and block, reflecting the fact that performance (as measured by proportional absolute error) was much worse in the first block of the first phase than in any other block. Thus, as had previously been found for time estimation, most of the learning occurred during the first block of training. In addition, as was predicted given specificity of training to the secondary task, there was also a significant three-way interaction of secondary-task continuity, phase, and block. This interaction shows that the improvement from the first block of training to the first block of testing was greater when the secondary-task conditions during training were the same as those during testing than when the secondary-task conditions were switched between training and testing (see Fig. 1, top panel). As will be discussed later, however, the specificity of training documented by the three-way interaction does not account for all of the learning that took place in this experiment, because performance in the switch group (not just that in the same group) was also worse on the first block of training than on the first block of testing, indicating the presence of some general learning as well as learning specific to the presence of the

Table 1 Overall analysis of variance for Experiment 1

Source	<i>df</i>	<i>F</i>	η^2	<i>p</i>
Between Subjects				
Estimation condition (E)	2	1.22	.017	.298
Secondary-task presence at test (T)	1	25.56	.151	<.001
Secondary-task continuity (C)	1	0.11	.001	.743
E × T	2	1.88	.026	.156
E × C	2	0.67	.009	.513
T × C	1	32.81	.186	<.001
E × T × C	2	0.27	.004	.762
S within groups (error)	144	(0.440)		
Within Subjects				
Phase (P)	1	34.45	.193	<.001
P × E	2	1.13	.015	.326
P × T	1	62.56	.303	<.001
P × C	1	12.19	.078	<.001
P × E × T	2	2.71	.036	.070
P × E × C	2	0.57	.008	.567
P × T × C	1	95.64	.399	<.001
P × E × T × C	2	1.74	.024	.180
P × S within groups (error)	144	(0.133)		
Block (B)	5	42.86	.229	<.001
B × E	10	1.27	.017	.245
B × T	5	1.30	.009	.262
B × C	5	0.36	.002	.877
B × E × T	10	1.50	.020	.135
B × E × C	10	0.97	.013	.471
B × T × C	5	0.36	.002	.874
B × E × T × C	10	1.31	.018	.220
B × S within groups (error)	720	(0.052)		
P × B	5	31.14	.178	<.001
P × B × E	10	1.63	.022	.095
P × B × T	5	1.17	.008	.320
P × B × C	5	2.84	.019	.015
P × B × E × T	10	0.86	.012	.574
P × B × E × C	10	0.60	.008	.817
P × B × T × C	5	0.63	.004	.677
P × B × E × T × C	10	0.92	.013	.510
P × B × S within groups (error)	720	(0.052)		

Values in parentheses represent mean squared errors. S = subjects

secondary task. In any event, the evidence for specificity of training was equivalent across the three estimation conditions; we found no hint of a four-way interaction of estimation condition, secondary task continuity, phase, and block.

Some components of the crucial significant three-way interaction (i.e., effects involving a subset of the three factors comprising the interaction) were also significant, including the main effect of block, the main effect of phase, and the interaction of secondary-task continuity and phase. Also

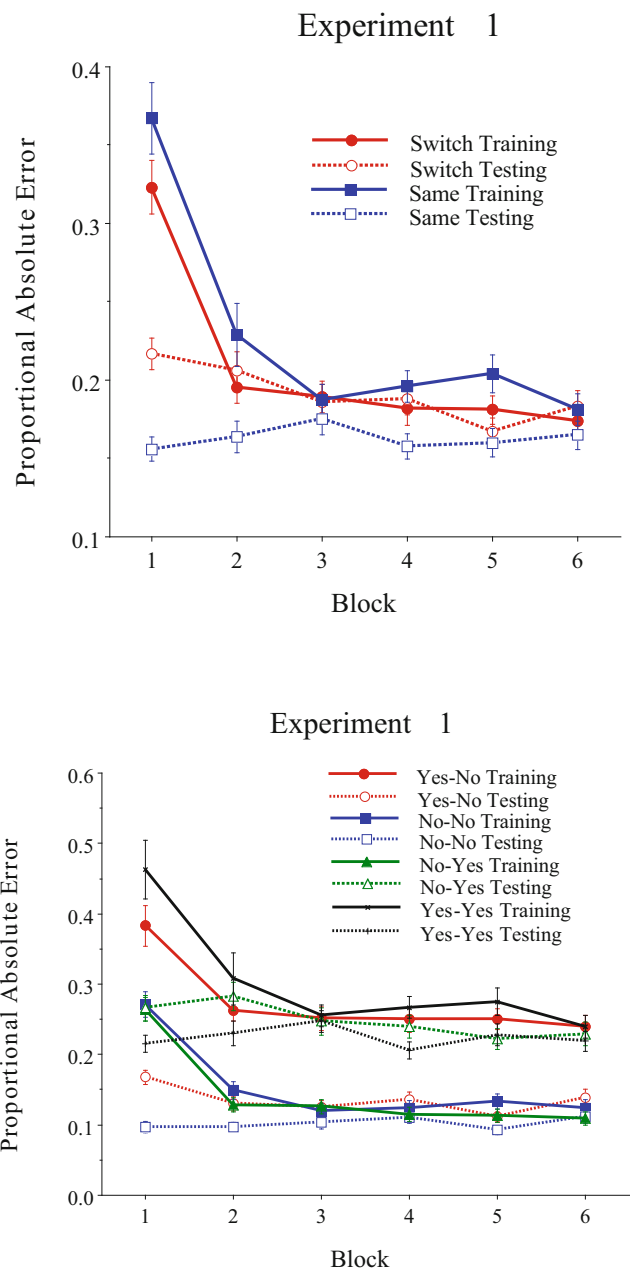


Fig. 1 Proportional absolute error in Experiment 1, as a function of secondary-task continuity (same, switch), phase, and block (top panel) and as a function of secondary-task presence at training (yes, no), secondary-task presence at test (yes, no), phase, and block (bottom panel). Error bars show between-subjects standard errors of the mean

significant was the interaction of secondary-task presence at test and phase, because naturally the secondary task's presence at test affected testing (no: $M = .118$, $SEM = .003$; yes: $M = .236$, $SEM = .005$) but not training (no: $M = .213$, $SEM = .005$; yes: $M = .222$, $SEM = .006$). The component main effect of secondary-task presence at test was also significant. In addition, we found a significant three-way interaction of secondary-task continuity, secondary-task presence at test, and phase, as well as the component two-way interaction of

secondary-task continuity and secondary-task presence at test. These interactions reflect the fact that the error magnitude was higher when there was a secondary task than when there was no secondary task in a given phase. Also, the advantage at testing for keeping the secondary task the same as at training, relative to switching, depended on the secondary-task presence at test: The advantage for the same over the switch group at testing was larger when there was no secondary task at test (training same: $M = .153$, $SEM = .005$; training switch: $M = .273$, $SEM = .008$; testing same: $M = .102$, $SEM = .003$; testing switch: $M = .135$, $SEM = .004$) than when there was a secondary task at test (training same: $M = .301$, $SEM = .011$; training switch: $M = .142$, $SEM = .005$; testing same: $M = .224$, $SEM = .006$; testing switch: $M = .248$, $SEM = .007$); see also Fig. 1, bottom panel.

Separate follow-up tests for each phase (training and testing) were conducted to explore further the crucial three-way interaction of secondary-task continuity, phase, and block. As we expected on the basis of specificity of training, for the testing phase the interaction of secondary-task continuity and block was significant, $F(5, 720) = 2.79$, $MSE = .036$, $p = .017$, $\eta^2 = .019$, whereas this same interaction was not significant for the training phase, $F(5, 720) < 1$.

Specificity and transfer To focus on specificity and transfer, a mixed factorial analysis of variance was conducted that was limited to the first block of training and the first block of testing. This analysis was partially redundant with the overall analysis, but it served to highlight the important patterns of results, especially because of the finding from the overall analysis that most of the learning occurred in the first block of training. This analysis also had the advantage of being similar to the Block 1 analysis used by Wohldmann et al. (2012) to examine specificity and transfer in the time estimation task. This analysis included the between-subjects factors Estimation Condition, Secondary-Task Continuity, and Secondary-Task Presence at Test, and the single within-subjects factor Phase. Evidence of specificity of training would take the form of an interaction between phase and secondary-task continuity, showing better performance during testing but not during training when the secondary-task condition was the same during training and testing, relative to when the secondary-task conditions were switched from training to testing. This crucial interaction was significant, $F(1, 144) = 9.61$, $MSE = .133$, $p = .002$, $\eta^2 = .063$, because we observed a lower proportional error for the same than for the switch group at testing, with a smaller difference (in the opposite direction) at training between same and switch (see Fig. 2, top panel), as we would expect, given specificity of training. Confirming this pattern, in separate tests of each phase, the main effect of secondary-task continuity was significant for the testing phase, $F(1, 144) = 15.63$, $MSE = .056$, $p < .001$, $\eta^2 = .098$, but not for the training phase, $F(1, 144) =$

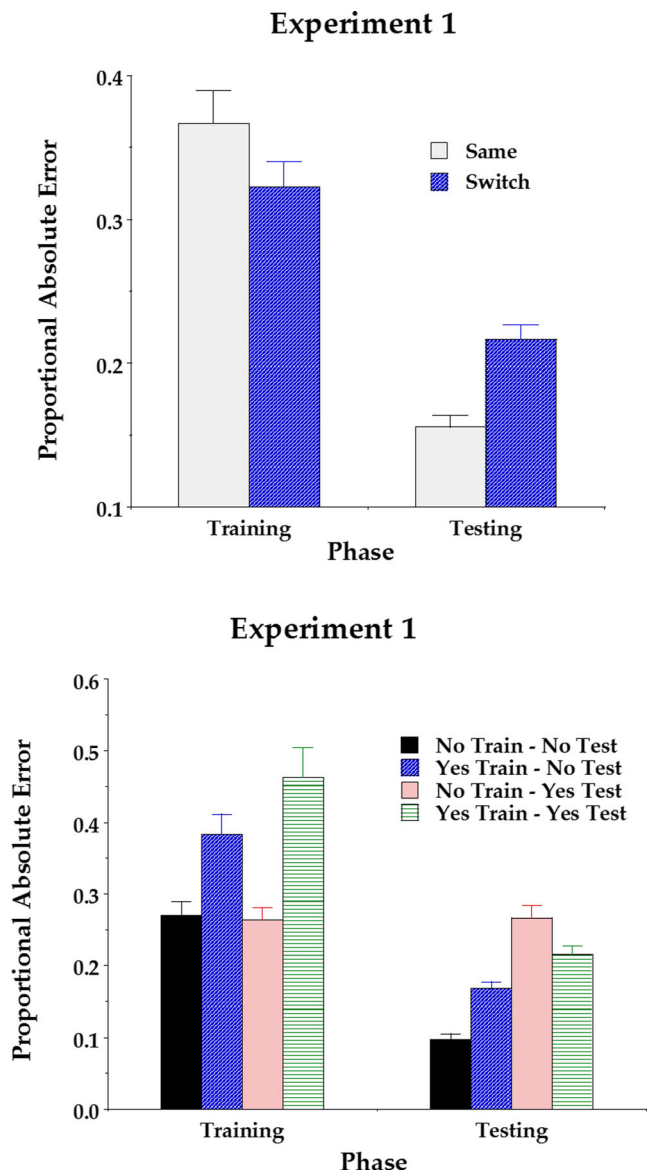


Fig. 2 Proportional absolute error in Block 1 of Experiment 1 for each phase, as a function of secondary-task continuity (same, switch) (top panel) and as a function of secondary-task presence at training (yes train, no train) and secondary-task presence at testing (yes test, no test) (bottom panel). Error bars show between-subjects standard errors of the mean

1.85 , $MSE = .239$, $p = .175$, $\eta^2 = .013$. The three-way interaction of phase, secondary-task continuity, and estimation condition was not significant, $F(2, 144) < 1$, indicating that specificity was observed for all three estimation conditions. However, as we mentioned in the consideration of the overall analysis, the observed specificity evident in Fig. 2 cannot explain all of the learning that occurred in this experiment: The difference between training and testing for the switch group is at least as large as the difference at testing between the same and switch groups, suggesting that learning has a general component, as well as a specific component due to the presence of the secondary task.

We found one significant interaction involving estimation condition—namely, a significant interaction of secondary-task presence at test and estimation condition, $F(2, 144) = 4.68$, $MSE = .162$, $p = .011$, $\eta^2 = .061$, reflecting the fact that the presence of the secondary task at test affected performance overall (averaged across the two phases) more in the time condition (no: $M = .223$, $SEM = .018$; yes: $M = .351$, $SEM = .029$) than in the distance condition (no: $M = .214$, $SEM = .013$; yes: $M = .308$, $SEM = .021$) or in the time line condition (no: $M = .253$, $SEM = .018$; yes: $M = .247$, $SEM = .013$), in which the means were actually reversed. In addition, there was a main effect of phase, $F(1, 144) = 88.47$, $MSE = .133$, $p < .001$, $\eta^2 = .381$, because of an overall reduction in proportional error from training ($M = .345$, $SEM = .014$) to testing ($M = .187$, $SEM = .007$). We also observed a significant three-way interaction of phase, secondary-task continuity, and secondary-task presence at test, $F(1, 144) = 18.65$, $MSE = .133$, $p < .001$, $\eta^2 = .115$, because the reduction in proportional error when the secondary tasks were the same at training and testing, relative to when they were switched between training and testing, was evident at testing both when the secondary task was present and not present at testing. In contrast, the reduction in proportional error for the same relative to the switch group was evident at training only when the secondary task was not present at testing. The difference between the same relative to the switch group at training was in the opposite direction when the secondary task was present at testing, presumably because of the higher proportional error when a secondary task was present, relative to when a secondary task was not present in a given phase (see Fig. 2, bottom panel). Furthermore, the component main effect of secondary-task presence at test, $F(1, 144) = 15.03$, $MSE = .162$, $p < .001$, $\eta^2 = .095$, the component interaction of secondary-task continuity and secondary-task presence at test, $F(1, 144) = 19.74$, $MSE = .162$, $p < .001$, $\eta^2 = .121$, and the component interaction of phase and secondary-task presence at test, $F(1, 144) = 4.44$, $MSE = .133$, $p = .037$, $\eta^2 = .030$, were all significant.

Discussion

In previous experiments of time estimation (e.g., Healy et al., 2005), test performance was better when secondary-task presence was the same at training and testing than when it was switched from training to testing, indicating that training of time estimation was specific to secondary-task presence. It was possible that this effect of specificity to the secondary task was unique for time estimation (or was an artifact of the time estimation procedure), because the secondary task filled the time interval being estimated. Nevertheless, the same training specificity of the primary task to the presence of a secondary task was evident in the present study for distance estimation as well as for time estimation with and without a growing line evident, and the magnitudes of the observed

specificity of training were statistically equivalent in the three conditions, despite another significant difference between them. Thus, the observed specificity cannot be attributed to the fact that the secondary task filled the time interval being estimated.

As we mentioned in the results, the learning specific to the secondary task was accompanied by more general learning, which was evident from the improvement from training to testing even when there was a switch in secondary-task presence.

Experiment 2

Experiment 1 verified that test performance on estimating intervals suffered when there was a switch from training to testing with respect to the presence of a secondary task during estimation, thereby demonstrating that training in estimation is specific to a secondary task that one engages in while estimating. The question arises whether specificity of training would also be evident if one switched between training to testing with respect to the FOP during interval estimation. To evaluate this question, in Experiment 2, subjects were trained in one estimation task (distance or time) and were then tested in either the same estimation task (i.e., with the same FOP) or the alternate estimation task (i.e., with the alternate FOP). In each case, subjects saw on the screen a line that grew at a fixed rate, with the horizontal starting position of the line varying randomly across trials. In this experiment, no secondary tasks were present during estimation in either primary task. Thus, the procedures for the subjects were identical for the two estimation tasks, with only the FOP (time or distance) varying across tasks.

Method

Subjects A total of 48 undergraduates from an introductory psychology course at the University of Colorado participated for course credit. They were divided into four subject groups, with 12 subjects in each group. Subjects were assigned to groups by a fixed rotation based on their time of arrival for testing.

Design The four subject groups varied as a function of training condition and task continuity. A mixed factorial design was, thus, used with two between-subjects variables—training condition (distance, time line) and task continuity (same, switch)—and two within-subjects variables—phase (training, testing) and block (1–6). The dependent variable was proportional absolute error.

Procedure The procedure was the same as in Experiment 1 for the distance and time line conditions, except that there was

no secondary task during either training or testing for any of the subjects.

Results

Overall analysis An initial overall mixed factorial analysis of variance was conducted to examine the acquisition and retention of the estimation skill across the six blocks of training and testing. The analysis included the between-subjects factors Training Condition and Task Continuity, along with the within-subjects factors Phase and Block. Of interest was the significant interaction of phase and block, $F(5, 220) = 18.53$, $MSE = .039$, $p < .001$, $\eta^2 = .296$, reflecting the fact that performance (as measured by proportional absolute error) was much worse in the first block of the first phase than in any other block. Thus, as in Experiment 1, most learning occurred during the first block of training. We also observed a significant main effect of block, $F(5, 220) = 14.55$, $MSE = .049$, $p < .001$, $\eta^2 = .249$, because of the overall improvement across blocks, and a significant three-way interaction of training condition, phase, and block, $F(5, 220) = 2.97$, $MSE = .039$, $p = .013$, $\eta^2 = .063$, because the two conditions did not have the exact same profile of performance across blocks, especially during testing (see Fig. 3). In separate analyses of the training and testing phases, the interaction of training condition and block was significant for testing, $F(5, 220) = 3.54$, $MSE = .007$, $p = .004$, $\eta^2 = .074$, but not for training, $F(5, 220) < 1$.

Specificity and transfer As in Experiment 1, to focus on training specificity and transfer, an analysis was limited to the first

block of training and the first block of testing. Training Condition, Task Continuity, and Phase were included as factors in a mixed factorial analysis of variance, with the first two factors (Training Condition and Task Continuity) varied between subjects and the third factor (Phase) varied within subjects. The only significant result was a main effect of phase, $F(1, 44) = 40.99$, $MSE = .016$, $p < .001$, $\eta^2 = .482$, reflecting overall improvement from training ($M = .316$, $SEM = .019$) to testing ($M = .151$, $SEM = .017$) in proportional error. Importantly, the improvements were equivalent when the tasks used in the two phases were the same and when they were switched; the interaction of phase and task continuity was not significant, $F(1, 44) < 1$. There was a trend for lower proportional error when the tasks were the same ($M = .212$, $SEM = .019$) than when they were switched ($M = .255$, $SEM = .024$); the main effect of task continuity was marginally significant, $F(1, 44) = 2.89$, $MSE = .016$, $p = .096$, $\eta^2 = .062$, but the effect of task continuity was at least as great during training (before any switch could occur) as during testing (after any switch in the task) (see Fig. 4). No significant effects involved training condition, $F(1, 44) < 1.53$, $p > .222$, in each case. These results indicate a clear pattern of “transfer,” with no indication of any training specificity.

Discussion

Estimation performance was better at test than at training for both time and distance estimation. Performance was no better at test when the training and testing conditions were the same than when they were switched, implying no specificity and perfect “transfer” of training to testing across estimation tasks.

Experiment 2

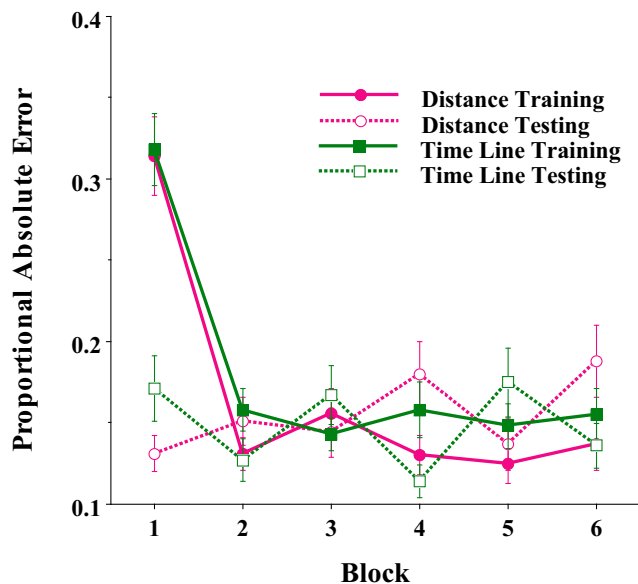


Fig. 3 Proportional absolute error in Experiment 2 as a function of training condition, phase, and block. Error bars show between-subjects standard errors of the mean

Experiment 2

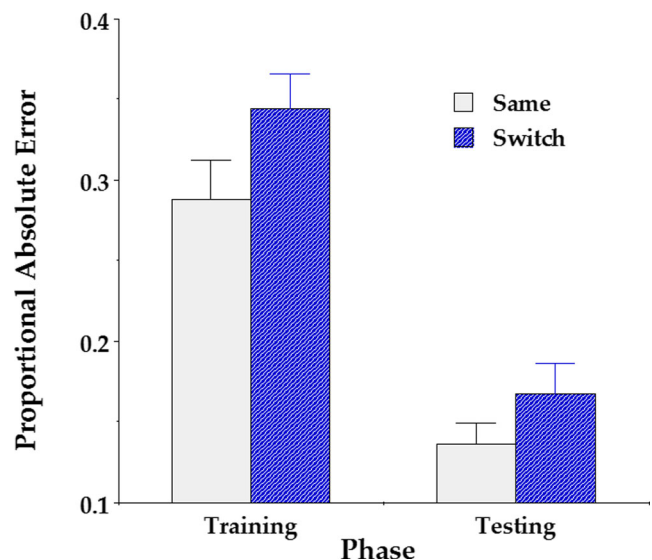


Fig. 4 Proportional absolute error in Block 1 of Experiment 2 as a function of task continuity (same, switch) and phase. Error bars show between-subjects standard errors of the mean

It is possible, however, that the lack of specificity was due to the presence of the growing line in both conditions, which might have influenced the subjects' estimation strategies even during time estimation, when there was no need to consider the line length. Experiment 3 addressed that issue by removing the growing line during the time estimation task.

Experiment 3

As in Experiment 2, subjects were trained in one estimation task (time or distance) and were then tested in either the same task or the alternate task. In the distance estimation condition, but not in the time estimation condition, subjects saw on the screen a line that grew at a fixed rate, with the horizontal starting position of the line varying randomly across trials.

Method

The method of Experiment 3 was identical to that of Experiment 2, except that the fixed-rate growing line was shown only in the distance condition, not in the time condition.

Results

Overall analysis Again, an initial overall mixed factorial analysis of variance was conducted to examine the acquisition and retention of the estimation skill across the six blocks of training and testing. The analysis included the between-subjects factors Training Condition and Task Continuity, along with the within-subjects factors Phase and Block. As in the previous experiments, of interest was the interaction of phase and block, $F(5, 220) = 5.28$, $MSE = .039$, $p < .001$, $\eta^2 = .107$, again showing that performance was much worse in the first block of the first phase than in any other block. Also, as previously, we observed a significant main effect of block, $F(5, 220) = 16.53$, $MSE = .007$, $p < .001$, $\eta^2 = .273$, because performance improved overall across blocks. In addition, and importantly, unlike in Experiment 2, there was a three-way interaction of task continuity, phase, and block, $F(5, 220) = 3.85$, $MSE = .007$, $p = .002$, $\eta^2 = .080$. This interaction shows that the improvement from the first block of training to the first block of testing held only when the training and testing blocks had the same FOP, not when they were switched (see Fig. 5), thus demonstrating clear specificity of training. In separate analyses of the training and testing phases, as we expected given specificity of training, the interaction of task continuity and block was significant for the testing phase, $F(5, 220) = 2.66$, $MSE = .006$, $p = .023$, $\eta^2 = .057$, but not for the training phase, $F(5, 220) = 1.62$, $MSE = .007$, $p = .157$, $\eta^2 = .035$.

Experiment 3

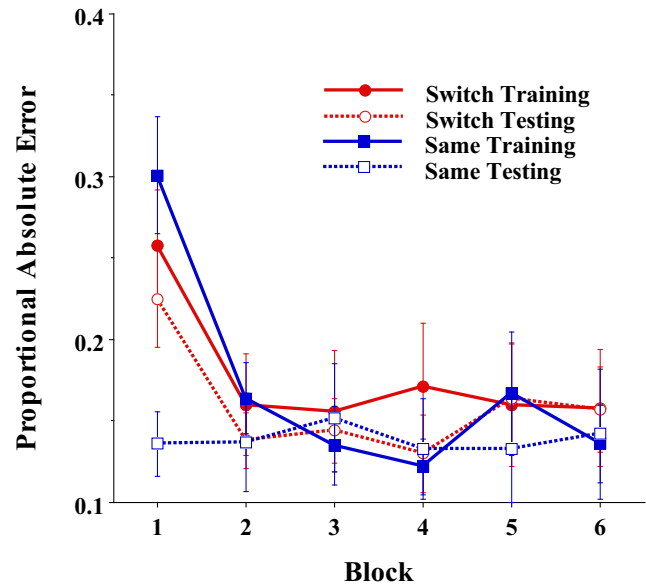


Fig. 5 Proportional absolute error in Experiment 3 as a function of task continuity (same, switch), phase, and block. Error bars show between-subjects standard errors of the mean

Specificity and transfer To focus on training specificity and transfer, as in Experiment 2, an analysis was limited to the first block of training and the first block of testing, including the factors Training Condition, Task Continuity, and Phase. As in Experiment 2, a main effect of phase emerged, $F(1, 44) = 17.97$, $MSE = .013$, $p < .001$, $\eta^2 = .290$, reflecting improvement from training ($M = .279$, $SEM = .025$) to testing ($M = .181$, $SEM = .019$) in proportional error. Importantly, the improvement was much larger when the tasks were the same at training and testing than when they were switched from training to testing (see Fig. 6); the interaction of phase and task continuity was significant, $F(1, 44) = 7.94$, $MSE = .013$, $p = .007$, $\eta^2 = .153$, but no other effects were: $F(1, 44) < 2.40$, $p > .128$, in each case. These results indicate a clear pattern of specificity of training, with essentially no indication of any transfer of training from one task to another. Confirming this pattern of specificity, in separate analyses of each phase, the main effect of task continuity was significant for the testing phase, $F(1, 44) = 6.18$, $MSE = .015$, $p = .017$, $\eta^2 = .123$, but not for the training phase, $F(1, 44) < 1$. In fact, a separate analysis restricted to the conditions involving a switch in estimation task from training to testing yielded no main effect of phase, $F(1, 22) < 1$, documenting the lack of transfer under those conditions, whereas a separate analysis restricted to the conditions involving the same estimation task in training and testing yielded a large effect of phase, $F(1, 22) = 26.96$, $MSE = .012$, $p < .001$, $\eta^2 = .551$, documenting robust improvement under those conditions. Thus, unlike in Experiment 1, we found no evidence for general learning beyond the learning specific to the estimation task in this experiment.

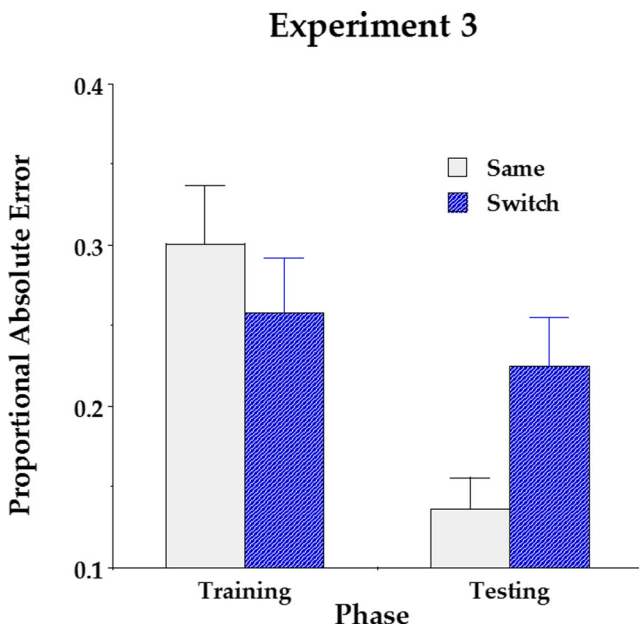


Fig. 6 Proportional absolute error in Block 1 of Experiment 3 as a function of task continuity (same, switch) and phase. Error bars show between-subjects standard errors of the mean

Discussion

The results for Experiment 3 were markedly different from those of Experiment 2 with respect to the question of any specificity of training. Experiment 2, which involved the external spatial cue (the growing line) for both the time and distance estimation tasks, yielded a clear pattern of “transfer” from one estimation task to another, with no evidence of specificity of training. In contrast, Experiment 3, which involved the external spatial cue for the distance estimation task but not for the time estimation task, yielded no evidence at all for transfer from one estimation task to another, indicating strong specificity of training in that case. An analysis of variance comparing the results from Experiments 2 and 3 for specificity and transfer assessed the magnitudes of these crucial differences. That analysis, like the earlier ones, was limited to the first block of training and the first block of testing, and included the factors Training Condition, Task Continuity, and Phase, as well as the factor Experiment. The analysis revealed a significant three-way interaction of experiment, phase, and task continuity, $F(1, 88) = 5.01$, $MSE = .014$, $p = .028$, $\eta^2 = .054$, documenting the different two-way interactions of phase and task continuity in the two experiments (cf. Figs. 4 and 6).

General discussion

In all three experiments, most of the learning occurred during the first block of the training phase, as was demonstrated by the much higher error rate in that block than in the other

blocks. Strong improvement was found as a result of training, as was evident both when comparing blocks in the training phase and when comparing the first block in training to the first block in testing, in all cases, with one important exception: No improvement was apparent from training to testing in Experiment 3, when the training and testing phases involved different estimation tasks (time vs. distance; i.e., different FOPs) and when a physical spatial cue (growing line) was present for the distance estimation task but not for the time estimation task. Specificity of training (better test performance when training and testing were the same than when they were switched) was evident in all cases, again with one important exception: There was no specificity of training in Experiment 2 when the only difference between training and testing was in the FOP (time or distance) and when a physical spatial cue (growing line) was present for both tasks. Specificity of training was even evident in Experiment 1, when the primary estimation task was held constant and only the presence of a secondary alphabet task varied between training and testing, and the amounts of specificity were statistically equal in magnitude for the three estimation tasks (time, distance, and time line). This specificity of training to the presence of a secondary task was, thus, not an artifact of the time estimation procedure, in which the secondary task filled the time interval being estimated, because it was also evident during the distance estimation procedure, when the secondary task occurred concurrently with the growth of the line length that was being estimated.

Perhaps specificity of training to the presence of a secondary task was evident in distance estimation because our subjects in distance estimation used the strategy of estimating time rather than distance when they were specifically instructed to estimate line length. However, the subjects must have relied on the growing line cue at least to some extent in the distance estimation task, and even to a degree in the time estimation task when the line cue was available, because the patterns of transfer and specificity across time and distance estimation tasks were markedly different in Experiment 2 (in which both tasks included the line cue, and there was “transfer” without specificity) and Experiment 3 (in which only the distance task included the line cue, and specificity of training was evident). The subjects’ attention to the growing line is understandable, given the well-established finding that attention is captured by the appearance of new perceptual objects (see, e.g., Yantis & Hillstrom, 1994).

In Experiment 2, the high “transfer” between time and distance estimation tasks suggests that training is not necessarily specific to the FOP when task-relevant external cues are the same between training and testing (i.e., when a line is present in all conditions). In contrast, in Experiment 3, the high specificity between training and testing suggests that training is specific to the FOP when the external task cues that potentially direct FOP differ between training and testing (i.e., a line is

present only in the distance conditions). Thus, in terms of the procedural reinstatement principle (e.g., Lohse & Healy, 2012), finding high specificity to the primary task in Experiment 3 but not in Experiment 2 implies that the procedures used for the primary tasks differed depending on the FOP, but only if the external task cues differed in the two FOP conditions.

It is possible that subjects in the time line condition used the exact same strategy as did subjects in the distance condition, because in both cases, as was just mentioned, they could rely on the external spatial cue—namely, the length of the fixed-rate growing line. Indeed, the lack of specificity across the time line and distance conditions in Experiment 2 is consistent with that possibility, and this suggests that no real “transfer” may have taken place in that experiment, because the two tasks might have been treated as equivalent by the subjects. However, this possibility that equivalent strategies were used by the subjects for the time line and distance estimation tasks does not diminish in any way the important result of Experiment 3, that the distance condition was treated quite differently from the time condition when a growing line was not present. There was strong evidence for specificity in that case, with no evidence of any general learning that transferred between the distance task and the time task without the external spatial cue.

In Experiment 1, we compared all three conditions and found no significant differences among them, in terms of either performance accuracy or the reliance of the primary task on the presence of a secondary task. Specificity of the primary task to the presence of a secondary task was found in each case, and to the same extent statistically. Thus, the present experiments together demonstrate that specificity to the presence of a secondary task is not unique to the task of time estimation, and occurs even for a task relying on spatial information, contrary to expectations based on previous demonstrations of transfer from one spatial task to another (e.g., Uttal et al., 2013). In terms of the procedural reinstatement principle, this finding of specificity to the presence of a secondary task in Experiment 1 implies that the procedures used for the primary task also differed depending on the presence of a secondary task.

As we mentioned in the Discussion of Experiment 1, some general learning was evident in that situation, as well as learning specific to the presence of the secondary task, because improvement took place from training to testing even when the secondary-task conditions were switched between the two phases. Because no general learning was evident in Experiment 3, which held constant for a given subject the secondary-task conditions, but such learning was evident in Experiment 1, which held constant for a given subject the presence of an external spatial cue, it seems likely that some learning components of time and distance estimation are specific to the presence of a secondary task, and that different

learning components are specific to the presence of external spatial cues.

How can the subjects make use of the secondary task in making their time or distance estimates? Wohldmann et al. (2010, 2012) provided evidence that when the time estimation task is coupled with a secondary alphabet task, which can be viewed as a graduated counting task, subjects use the number of letters recited as a way to keep track of the amount of time that has elapsed, thereby combining the primary time estimation task with the secondary alphabet task into a single functional task. Such a strategy would also be possible for the distance estimation task used here, because the line cue grew at a constant rate. In future research, it would be worthwhile to determine whether a different pattern of results would occur with distance estimation if the line were to grow instead at a variable rate, so that time could no longer be a reliable cue for distance. In that case, the number of letters recited would not be a valid index of line distance, because a given line distance would be independent of how quickly or slowly it was achieved. Indeed, in such a case, time estimation would not be a valid proxy for distance estimation, so the two tasks would differ not just in FOP, as they did in the present study. Under such circumstances, transfer from one estimation task (time or distance) to the alternate task would be less likely, even when the two tasks included the same external spatial cues.

In any event, the most important lesson learned from the present study is that the high degree of training specificity found for the time estimation task can also be found for a parallel task involving distance estimation. Subjects performed better on the estimation task during testing when the presence of a secondary task was the same at training and testing than when the secondary task switched, even when no changes were made across training and testing in the primary estimation task. Furthermore, although the two tasks of time and distance estimation showed perfect “transfer” when they differed only in their FOPs and had available the same external spatial cues, they showed no transfer when the external spatial cues were present in the distance task but not in the time task. Such striking findings of training specificity should serve as a counterpoint to recent claims that training in one task aids performance in a very different task (see, e.g., Uttal et al., 2013, for spatial tasks, but see also Shipstead et al., 2012, for a review of recent claims for working memory tasks and a similar caution about them). The present findings imply, instead, that for a training program to be most effective, it needs to anticipate all aspects of the primary task, and even the secondary-task requirements, that will occur during eventual testing.

Author note This research was supported in part by National Aeronautics and Space Administration Grant Numbers NNX10AC87A and NNX14AB75A to the University of Colorado. We thank the members

of the Center for Research on Training at the University of Colorado for helpful suggestions about this research. We are indebted to Robert Proctor, Timothy Rickard, Steven Pan, and an anonymous reviewer for helpful comments on an earlier version of this article; Mary Guhl and Simone Addison for help testing the subjects in Experiment 1; Lakshmi Lalchandani for help with data tabulation; Mark Thrasher for help testing the subjects in Experiments 2 and 3; and Ernest Mross for programming the experiments. Experiments 2 and 3 were summarized in a poster at the 2011 meeting of the Psychonomic Society in Seattle, WA. Experiment 1 was conducted after Experiments 2 and 3, but is reported first for ease of exposition.

References

- Clawson, D. M., King, C. L., Healy, A. F., & Ericsson, K. A. (1995). Training and retention of the classic Stroop task: Specificity of practice effects. In A. F. Healy & L. E. Bourne Jr. (Eds.), *Learning and memory of knowledge and skills: Durability and specificity* (pp. 234–254). Thousand Oaks, CA: Sage.
- Healy, A. F. (1975). Coding of temporal-spatial patterns in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*, 481–495. doi:10.1016/S0022-5371(75)80026-0
- Healy, A. F., Wohldmann, E. L., & Bourne, L. E., Jr. (2011). How does practice with a reversed mouse influence subsequent speeded aiming performance? A test of global inhibition. *Journal of Cognitive Psychology*, *23*, 559–573. doi:10.1080/20445911.2011.547467
- Healy, A. F., Wohldmann, E. L., Parker, J. T., & Bourne, L. E., Jr. (2005). Skill training, retention, and transfer: The effects of a concurrent secondary task. *Memory & Cognition*, *33*, 1457–1471. doi:10.3758/BF03193378
- Healy, A. F., Wohldmann, E. L., Sutton, E. M., & Bourne, L. E., Jr. (2006). Specificity effects in training and transfer of speeded responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 534–546. doi:10.1037/0278-7393.32.3.534
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the USA*, *105*, 6829–6833. doi:10.1073/pnas.0801268105
- Kolers, P. A., & Roediger, H. L., III. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior*, *23*, 425–449. doi:10.1016/S0022-5371(84)90282-2
- Lohse, K. R., & Healy, A. F. (2012). Exploring the contributions of declarative and procedural information to training: A test of the procedural reinstatement principle. *Journal of Applied Research in Memory and Cognition*, *1*, 65–72. doi:10.1016/j.jarmac.2012.02.002
- McDaniel, M. A., Friedman, A., & Bourne, L. E., Jr. (1978). Remembering the levels of information in words. *Memory & Cognition*, *6*, 156–164. doi:10.3758/BF03197441
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*, 519–533. doi:10.1016/S0022-5371(77)80016-9
- Rickard, T. C., & Bourne, L. E., Jr. (1996). Some tests of an identical elements model of basic arithmetic skills. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1281–1295. doi:10.1037/0278-7393.22.5.1281
- Rickard, T. C., Healy, A. F., & Bourne, L. E., Jr. (1994). On the cognitive structure of basic arithmetic skills: Operation, order, and symbol transfer effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 1139–1153. doi:10.1037/0278-7393.20.5.1139
- Roediger, H. L., III, Weldon, M. S., & Challis, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3–41). Hillsdale, NJ: Erlbaum.
- Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*, 628–654. doi:10.1037/a0027473
- Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.
- Sternberg, R. J. (2008). Increasing fluid intelligence is possible after all. *Proceedings of the National Academy of Sciences of the USA*, *105*, 6791–6792. doi:10.1073/pnas.0803396105
- Thorndike, E. L. (1906). *The principles of teaching: Based on psychology*. New York, NY: A. G. Seiler. doi:10.1037/11487-000
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352–373. doi:10.1037/h0020071
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, *139*, 352–402. doi:10.1037/a0028446
- Wohldmann, E. L., Healy, A. F., & Bourne, L. E., Jr. (2010). Task integration in time production. *Attention, Perception, & Psychophysics*, *72*, 1130–1143. doi:10.3758/APP.72.4.1130
- Wohldmann, E. L., Healy, A. F., & Bourne, L. E., Jr. (2012). Specificity and transfer effects in time production skill: Examining the role of attention. *Attention, Perception, & Psychophysics*, *74*, 766–778. doi:10.3758/s13414-012-0272-5
- Yantis, S., & Hillstrom, A. P. (1994). Stimulus-driven attentional capture: Evidence from equiluminant visual objects. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 95–107. doi:10.1037/0096-1523.20.1.95