Children’s Task Performance under Stress and Non-stress Conditions:

A Test of the Processing Efficiency Theory

EeLynn Ng and Kerry Lee

National Institute of Education, Singapore

Correspondence concerning this article should be addressed to EeLynn Ng, Centre for Research in Pedagogy and Practice, National Institute of Education, 1 Nanyang Walk, Singapore 637616. Telephone: (65) 9001 4172, Fax: (65) 6896 9410. E-mail: eelynn.ng@gmail.com
Abstract

The effects of test anxiety on children’s task performance are not well-understood. We examined this issue using the processing efficiency theory (PET; Eysenck & Calvo, 1992) as a theoretical framework. High and low trait test-anxious children (N = 90) performed a mental arithmetic task under high and low memory load conditions. Each child performed the task under stressful and non-stressful conditions. Measures of task accuracy and reaction time served as indicators of performance effectiveness and processing efficiency, respectively. Consistent with the PET, processing efficiency, but not performance effectiveness, was detrimentally affected by test anxiety. However, we did not find support for the PET’s assumption that state anxiety mediates the test anxiety-task performance relationship. The roles of task demands on WM capacity and individual differences in WM capacity as moderators of the relationship between test anxiety and task performance will also be discussed.
Children’s Task Performance under Stress and Non-stress Conditions:

A Test of the Processing Efficiency Theory

This study contributes to a growing literature that investigates the relationship between anxiety and academic underachievement (Johnson & Gronlund, 2009; Owens, Stevenson, Norgate, & Hadwin, 2008). In particular, we focus on the cognitive consequences of test anxiety on mathematical performance. The processing efficiency theory (PET), both in its original (Eysenck & Calvo, 1992) and revised version (i.e., the attentional control theory, ACT; Eysenck, Derakshan, Santos, & Calvo, 2007) asserts that working memory (WM) mediates the effects of test anxiety on task performance. Although aspects of this prediction have found support in several studies (Hadwin, Brogan, & Stevenson, 2005; Ikeda, Iwanaga, & Seiwa, 1996), most of them were conducted with undergraduates (McDonald, 2001). Because children are likely to have less expertise in taking examinations but can be just as much if not more affected by poor examination performance, we examined the effects of test anxiety on children.

One issue on which we focused was the effects of trait versus state anxiety (Zeidner, 1998) on task performance. Trait test anxiety (trait TA) refers to a more stable aspect of an individual’s disposition to experience anxiety in evaluative situations. State TA refers to an individual’s test anxiety in reaction to a present stressor; it is determined interactively by trait anxiety and situational stress (Eysenck & Calvo, 1992). The PET assumes that state anxiety is responsible for anxiety-related performance impairments. This assumption has not been adequately addressed as few studies include a situational stressor to compare performance under high and low state anxiety. In some studies that included this manipulation, (e.g., Sorg & Whitney, 1992), different types of tasks were used in the stressful and non-stressful conditions. As a result, performance impairments cannot be unambiguously attributed to differences in state
anxiety. Others examined differences between high and low trait anxiety under stressful, but not non-stressful conditions (e.g., Ikeda et al., 1996).

Our design included trait TA and situational stress to assess their joint effects on task performance. We adapted Prins, Groot, and Hanewald’s (1994) procedure, in which children performed a mathematics test under high and low threat conditions. In the high threat condition, children were told that their test scores were important for their final class grade. In the other condition, children were told that test scores would not affect their final grade. We measured changes in state TA using a self-report measure which assessed subjective, consciously perceived feelings of apprehension and worry (Spielberger, 1973). Unlike biological measures of anxiety (e.g., skin conductance and heart rate), a subjective measure focuses on the cognitive component of anxiety, which is of primary concern in our study.

A second issue addressed in this study was the role of WM capacity in the relationship between test anxiety and task performance. This issue was addressed from both an experimental and a differential perspective. From an experimental perspective, Ashcraft and Kirk (2001) showed that availability of WM resources moderated the relationship between mathematics anxiety and task performance. We extended this investigation by testing the PET’s prediction that the adverse effects of test anxiety become greater as task demands on WM resources increase. MacLeod and Donnellan (1993) argued that this issue was not properly addressed in previous studies as the tasks’ extraneous characteristics were inadequately controlled. This criticism also applies to more recent studies. In Hadwin et al.’s (2005) study, participants performed the forward digit span, backward digit span and spatial WM tasks. In addition to likely differences in the amount of WM resources required to perform each task, the tasks also used different types of stimuli (i.e., pictures vs. numbers), which may have different demands on
WM resources (Salway & Logie, 1995). To address this limitation, we used MacLeod and Donnellan’s (1993) loading paradigm, which allows WM demands to be varied without using different task procedures at different memory load levels. We used a mental arithmetic task and a memory load task. The arithmetic problems consisted of two 2-digit carry problems (e.g., 75 + 29) presented horizontally. Problems of this type were selected as the problem-solving process is known to rely on WM resources (Ashcraft & Kirk, 2001; Trbovich & LeFevre, 2003).

From a differential perspective, we investigated whether individual differences in WM capacity moderated the test anxiety-task performance relationship. In Seyler, Kirk, and Ashcraft’s (2003) study, undergraduates performed a letter recall task and a demanding subtraction task simultaneously. Low WM span individuals demonstrated a larger decrease in recall accuracy than medium and high span individuals. Their findings suggest that although anxiety results in a general reduction in individuals’ WM resources, having a higher WM capacity should serve as a buffer against this effect. In support of this hypothesis, Johnson and Gronlund (2009) found trait anxiety correlated negatively with task accuracy at low and average levels of WM capacity, but not at high levels of WM capacity.

Summary

This study compared the performance of high and low trait TA groups on an experimental task with varying WM demands. Both groups performed the task under evaluative and non-evaluative task conditions. These task conditions were used to introduce high versus low situational stress, respectively. In the evaluative condition, the high trait TA group was expected to report a larger increase in state TA from pre-test to post-test compared to the low trait group. In the non-evaluative condition, both trait groups were expected to report smaller increases in state TA from pre- to post-test.
Following the PET’s distinction between performance effectiveness and processing efficiency, these were assessed by accuracy and reaction time (RT) measures, respectively. RT measures are commonly used to compare the performances of high- and low-anxious individuals (Eysenck et al., 2007). In line with the PET, both trait TA groups should perform equally accurately regardless of task condition. Task accuracy was expected to decrease from low to high memory load because of the added difficulty of performing two demanding tasks simultaneously under high load. Regarding RT measures, our predictions differed from previous studies (e.g., MacLeod & Donnellan, 1993), which predicted trait group differences only on the primary task. We expected anxiety-related increases in RT on both the mental arithmetic and memory load tasks because we instructed our participants to pay equal attention to both. Specifically, we expected the RT measures to be affected by a 3-way interaction. In the evaluative condition, the high trait TA group should show a larger increase in RT as memory load increases compared to the low trait TA group. In the non-evaluative condition, both groups should show a similar magnitude of increase in RT as memory load increases. To determine the impact of individual differences on these predictions, we examined the relationship between test anxiety and task performance after controlling for individual differences in WM capacity.

Method

Participants and Design

A total of 114 10 year-old pupils from three public schools in Singapore participated with parental consent. We used the Test Anxiety Inventory (TAI; Spielberger et al., 1980) to categorize participants into high and low trait TA groups. Since the TAI does not provide normed scores for children, percentile splits of the TAI scores (> 70th percentile vs. < 40th percentile) were used to classify participants into high ($n = 45$, $M_{TAI} = 55.53$, $SD_{TAI} = 5.33$,}
Children’s Task Performance

$M_{\text{age}}= 10.86 \text{ years}, SD_{\text{age}} = 0.38 \text{ years}$ and low ($n = 45$, $M_{\text{TAI}} = 34.51$, $SD_{\text{TAI}} = 4.48$, $M_{\text{age}} = 10.79 \text{ years}, SD_{\text{age}} = 0.63 \text{ years}$) trait TA groups. Participants whose scores did not fall within the prescribed limits were excluded. A $t$-test confirmed significant differences in the TAI between the two groups, $t(88) = -20.25, p < .05.$

This study used a $2$ (trait TA: high vs. low) x $2$ (memory load: high vs. low) x $2$ (task condition: evaluative vs. non-evaluative) split-plot design. Memory load and task condition were manipulated on a within-subjects basis. WM capacity served as the covariate. The dependent measures were accuracy and RT. Only trials on which participants were accurate on both the mental arithmetic and memory load tasks were deemed correct. The joint accuracy measure ensures comparability with other studies (e.g., Ashcraft & Kirk, 2001; Trbovich & LeFevre, 2003) and provides a more accurate gauge of performance on trials in which participants complied with task instructions, which emphasized the importance of paying equal attention to both tasks. Individual task accuracy measures were retained only for examining possible speed-accuracy trade-offs in task performance. RT measures were based on jointly accurate trials only.

Materials

Trait test anxiety. The TAI consisted of 20 statements describing various reactions towards tests and examinations. Pupils rated how they generally felt about each statement on a scale of 1 (Almost never) to 4 (Almost always). TAI scores range from 20 to 80.

State test anxiety. The State version of the State-Trait Anxiety Inventory for Children (STAIC; Spielberger, 1973) was administered immediately before and after participants had completed the experimental task. In the pre-test assessment, participants chose words that described how they felt at the present moment. In the post-test assessment, participants
responded according to how they felt while performing the task. STAIC scores range from 20 to 60.

*Experimental task.* Each trial consisted of a mental arithmetic task and a memory load task. The mental arithmetic task consisted of addition and subtraction problems. The memory load task consisted of a contiguous alphabet sequence in the low load condition (e.g., “A B C D E F”) and similar sequences presented in randomized order in the high load condition (e.g., “D A E C B F”). The task was presented on a computer using E-Prime (Schneider, Eschman, & Zuccolotto, 2002). The experimental task was preceded by two practice blocks of 16 trials each. Experimental trials were presented in four blocks of 16 trials each. The first and third blocks contained addition problems, while the remainder consisted of subtraction problems. Details of the task are shown in Figure 1.

Working memory (WM) capacity. Three central executive subtests (Listening Recall, Counting Recall and Backward Digit Recall) from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001) were administered. Each task was presented using a span approach. For example, in the Counting Recall task, participants counted arrays of dots on a series of cards. Each trial contained a series of cards, ranging from one to seven cards. At the end of each trial, participants recalled the number of dots on each card in the order of presentation. Practice trials were given prior to the actual trials to determine each participant’s starting span. The dependent measure for all tests was the number of trials correctly recalled.
Procedure

All tasks and measures were administered on a one-on-one basis over three sessions. During the first session, the WM capacity measure was administered first, followed by the trait TA measure and a practice version of the experimental task. The remaining sessions involved administration of the experimental task under evaluative and non-evaluative task conditions. The two task conditions were administered two days apart and the order of presentation of the two conditions was counterbalanced across participants.

In both task conditions, participants completed a pre-test measure of state TA, followed by the experimental task and a post-test measure of state TA. Task instructions for the evaluative and non-evaluative conditions were given upon completion of the pre-test measure. In the evaluative condition, participants were told that they will be tested on their mental arithmetic skills and that their results will be recorded in their school report books if they failed the test. In the non-evaluative condition, participants were told that the experimenter was interested in the length of time required to complete the task. Instructions for both conditions emphasized the importance of maintaining a high level of accuracy in the mental arithmetic and memory load tasks.

Results

Data Screening

Distributions of all variables were screened for univariate outliers and deviations from normality. Screening procedures were conducted separately for low and high trait TA groups. For state TA, WM capacity and joint task accuracy measures, scores larger than three SDs from their respective means were replaced by values at three SDs. This procedure ensured that the extreme scores could be retained without them having an adverse effect on the distribution of
scores. Replacement of univariate outliers rendered the distributions of all variables approximately normal.

The RT measures were subjected to a two-stage trimming procedure. The first stage involved removal of trials that exceeded RT thresholds, which were determined by examination of overall raw score distributions for each experimental condition separately. RT thresholds for the mental arithmetic task are 22500 ms for evaluative-low memory load (E-Low) and non-evaluative-low memory load (NE-Low) conditions, and 25000 ms for E-High and NE-High conditions. RT thresholds for the memory load task are 7000 ms for E-Low and NE-Low conditions, 9000 ms for E-High condition and 8000 ms for NE-High condition. No more than 5.0% of the trials were affected by this procedure. At the second stage, each participant’s trial-by-trial RT distribution was screened using the same procedure that was used for the state TA measure. Following this procedure, each participant’s mean RT on the mental arithmetic and memory load tasks were calculated for the four experimental conditions separately. All the mean RT distributions had skewness values below one after replacement of univariate outliers. Means and standard deviations for all variables are in Table 1.

State TA Scores

State TA scores were subjected to a 2 (trait TA: high vs. low) x 2 (task condition: evaluative vs. non-evaluative) x 2 (time: pre-test vs. post-test) split-plot ANOVA with trait TA as a between-subjects factor. The main effect of trait TA was significant, $F(1, 88) = 14.49, p < .01$, partial $\eta^2 = .14$. The high trait TA group reported higher state TA scores than the low trait
TA group (means of 36.55 and 31.81, respectively). The main effect of time, $F(1, 88) = 27.81, p < .01$, partial $\eta^2 = .24$ was also significant. State TA at post-test ($M = 36.09$) was higher than state TA at pre-test ($M = 32.27$). All 2-way interactions were not significant. Notably, the 3-way interaction was also not significant.

**Accuracy and RT Measures on Experimental Task**

Analyses of separate accuracy and RT measures revealed no significant evidence of speed-accuracy trade-offs on the experimental task. A 2 (trait TA: high vs. low) x 2 (memory load: high vs. low) x 2 (task condition: evaluative vs. non-evaluative) split-plot ANOVA was conducted for the joint task accuracy measure. A significant interaction between memory load and task condition, $F(1, 88) = 5.39, p < .05$, partial $\eta^2 = .06$, was found. The decline in accuracy from low to high memory load was greater in the evaluative task condition (low load: $M = 27.13$; high load: $M = 20.51$) compared to the non-evaluative task condition (low load: $M = 26.44$; high load: $M = 21.11$).

The mental arithmetic and memory load task RT measures were subjected to a 2 (trait TA: high vs. low) x 2 (memory load: high vs. low) x 2 (task condition: evaluative vs. non-evaluative) split-plot MANOVA. The 3-way interaction was significant, $F(2, 87) = 3.57, p < .05$, partial $\eta^2 = .08$. The interaction between memory load and task condition, $F(2, 87) = 4.07, p < .05$, partial $\eta^2 = .09$, and the main effect of memory load, $F(2, 87) = 137.09, p < .01$, partial $\eta^2 = .76$, were also significant.

To follow up on the MANOVA findings, separate 2 x 2 x 2 univariate ANOVAs were conducted for each RT measure. Mental arithmetic task RT was significantly affected by memory load, $F(1, 88) = 16.01, p < .01$, partial $\eta^2 = .15$. Participants responded more quickly in the low memory load condition ($M = 8827$ ms) than in the high load condition ($M = 9270$ ms).
For memory load task RT, the interaction between memory load and task condition, $F(1, 88) = 7.62, p < .05$, partial $\eta^2 = .08$, and the main effect of memory load, $F(1, 88) = 276.08, p < .01$, partial $\eta^2 = .76$, were significant. These effects were qualified by a significant 3-way interaction, $F(1, 88) = 7.21, p < .05$, partial $\eta^2 = .08$. To determine the nature of this interaction, separate 2 (memory load: low vs. high) x 2 (task condition: evaluative vs. non-evaluative) within-subjects ANOVAs were conducted for the high and low trait TA groups (with Bonferroni’s correction for alpha levels). For the low trait TA group, a significant main effect of memory load was found. Responses were faster in the low load condition than in the high load condition, $t(44) = -13.58, p < .01$. A significant 2-way interaction was observed for the high trait TA group, where the increase in RT from low to high memory load was 727 ms in the evaluative condition, $t(44) = -9.33, p < .01$, compared to 397 ms in the non-evaluative condition, $t(44) = -5.68, p < .01$.

A series of ANCOVAs and MANCOVAs were conducted to determine whether the relationship between test anxiety and task performance depended on individual differences in WM capacity. For the joint task accuracy measure, the ANCOVA and ANOVA findings were similar. The crucial 3-way interaction for the RT measures remained significant in the MANCOVA.

Discussion

Performance under Evaluative and Non-evaluative Conditions

Consistent with the PET, we showed that children’s processing efficiency, but not performance effectiveness, was detrimentally affected by test anxiety. Both trait TA groups exhibited the same pattern of performance effectiveness on the experimental task. In terms of processing efficiency, trait TA group differences were observed on the memory load RT. As memory load increased, the low trait TA group showed the same magnitude of reduction in
processing efficiency in the evaluative and non-evaluative conditions. In contrast, the high trait TA group showed a smaller reduction in processing efficiency in the non-evaluative condition, compared to both their own performance in the evaluative condition and the low trait TA group’s performance in the non-evaluative condition.

Effects of the Situational Stress Manipulation Procedure

Findings from the state TA measure show that the manipulation procedure did not significantly affect state TA. As our main findings concern the manipulation procedure, this finding has important implications for our interpretation. Although the state TA findings indicate that the manipulation did not work, findings from the performance measures suggest otherwise. Specifically, the 3-way interaction for memory load RT indicates that the manipulation had an effect on memory load task performance. One possibility is that the state TA manipulation did not have a large effect on participants’ perception of their state TA levels. However, it affected participants’ problem-solving behaviors, namely processing efficiency on the memory load task. This is reflected in the different patterns of performance shown by the high and low trait TA groups on memory load RT. Similar to our findings, Keogh and French (1997) failed to find a significant interaction between mood manipulation and occasion of testing on state anxiety scores but they observed a significant interaction between mood manipulation and trait anxiety group on task RT. The time at which the state TA measure was administered may also have affected its sensitivity. Participants rated their “at-task” (i.e., how they felt during the task) state TA levels after completing the experimental task. Due to the delay in the measurement of state TA, these ratings may not accurately reflect their emotional state during the task. In comparison, RT measures may have served as a more sensitive, real-time assessment of the effects of state TA on task performance. Ultimately, our data clearly show that the manipulation procedure did
affect memory load-related task performance. However, whether this effect is mediated by state TA remains an open question because of the nonsignificant effect of the manipulation on state TA.

Role of WM Capacity

We noted two findings which were inconsistent with the PET’s prediction. First, as memory load increased in the evaluative condition, the high trait TA group did not show a significantly larger reduction in processing efficiency than the low trait group. Compared to their adult counterparts’ performance under similar task conditions (e.g., Ashcraft & Kirk, 2001; MacLeod & Donnellan, 1993), our high trait TA participants performed the task more efficiently. Consistent with the PET, we postulate that high trait TA individuals are particularly concerned about their test performance, which resulted in a stronger motivation perform well on the task. To this end, children strived to complete the task as quickly as possible as a time limit was imposed in the evaluative condition. A possible reason for the divergence between our findings and what would be expected from the PET is that there is a developmental decline in academic intrinsic motivation (Gottfried, Marcoulides, Gottfried, Oliver, & Guerin, 2007). This decline results in a weaker influence of the motivational factor on adults’ performance compared to children. As a consequence, the difference in task efficiency between evaluative and non-evaluative conditions is greater for adults than it is for children.

A second finding that is inconsistent with the PET is that the high trait TA group showed a larger mean RT than the low trait group at low memory load in the non-evaluative condition. We propose that this finding may also be related to the academic intrinsic motivational factor. Because the non-evaluative condition was not a test and did not impose a time limit, the high
trait TA group may have been less motivated to perform well on the task. As a consequence, they were less efficient in completing the task.

From a differential perspective, inclusion of the covariate (i.e., WM capacity) reduced the sum of squared errors across all task performance measures, but it did not provide a significant adjustment to the marginal means of the two trait TA groups. Since the pattern of task performance did not change after considering the covariate’s effect, this suggests that individual differences in WM capacity did not have a strong influence on task performance. It may be the case that our task was not overly demanding of participants’ WM resources. Therefore, they were able to bring in additional processing resources in order to complete the task, regardless of each individual’s enduring WM capacity.

*Test Anxiety’s Effects on Primary and Secondary Tasks*

The findings differ from our hypothesis that trait group differences in processing efficiency would be found on both tasks. Previous studies also reported mixed findings. MacLeod and Donnellan (1993) reported trait anxiety group differences in efficiency on the grammatical reasoning task (the equivalent of which is our mental arithmetic task). In contrast, Derakshan and Eysenck (1998) found group differences on the memory task and the grammatical reasoning task. One explanation for this inconsistency is that the locus of the test anxiety effect depends on the specific experimental setup. In our study, the memory load stimuli must be committed to memory whereas the mental arithmetic stimuli remain visible until an answer is produced. Furthermore, maintenance of the memory load stimuli is susceptible to interference from the mental arithmetic stimuli. Thus, memory load task performance is highly dependent on attentional control processes to retain information while processing incoming stimuli as well as to inhibit interfering information. According to the ACT (Eysenck et al., 2007), attentional
control processes are susceptible to the negative effects of anxiety. Therefore, it is not surprising that trait TA group differences in processing efficiency were found on the memory load task.

Conclusions

We replicated an established finding in the literature by showing that test anxiety adversely affects children’s processing efficiency, but not performance effectiveness. Specifically, trait TA group differences were observed on memory load task efficiency. However, it is unclear whether this effect was mediated by state TA. Our findings were inconsistent with the PET’s prediction that the adverse effects of test anxiety increases as the task’s WM demands increase. Further, individual differences in WM capacity did not moderate the relationship between test anxiety and task performance. Possible explanations were offered to account for these findings. However, these are speculative at best and require further investigation in future studies.
References


Johnson, D. R., & Gronlund, S. D. (2009). Individuals lower in working memory capacity are particularly vulnerable to anxiety’s disruptive effect on performance. *Anxiety, Stress and Coping, 22*, 201-213.


Author Note

EeLynn Ng and Kerry Lee, Centre for Research in Pedagogy and Practice, National Institute of Education, Singapore.

The research presented in this paper is based on the first author’s doctoral dissertation. Earlier analyses of the data were presented at the biennial meeting of the International Society for the Study of Behavioral Development in July 2008. This research was supported by grants from the Centre for Research in Pedagogy and Practice and the Pradap Kow Foundation.
Table 1

Means and Standard Deviations of Study Variables for Low Trait Test Anxiety (TA; n = 45) and High Trait TA (n = 45) Groups

<table>
<thead>
<tr>
<th></th>
<th>State Anxiety(^a)</th>
<th>Task Accuracy(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Evaluative</td>
<td>Non-evaluative</td>
</tr>
<tr>
<td></td>
<td>Pre-Test</td>
<td>Post-Test</td>
</tr>
<tr>
<td>Low Trait TA</td>
<td>30.26 (5.77)</td>
<td>33.89 (7.92)</td>
</tr>
<tr>
<td>High Trait TA</td>
<td>34.21 (7.04)</td>
<td>39.89 (8.69)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Mental Arithmetic Task RT (msec)</th>
<th>Memory Load Task RT (msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Evaluative</td>
<td>Non-evaluative</td>
</tr>
<tr>
<td></td>
<td>Low Load</td>
<td>High Load</td>
</tr>
<tr>
<td>Low Trait TA</td>
<td>8355.23 (2894.09)</td>
<td>8851.68 (3609.55)</td>
</tr>
<tr>
<td>High Trait TA</td>
<td>9391.29 (2946.63)</td>
<td>9665.36 (3503.56)</td>
</tr>
</tbody>
</table>

Note. Values in parentheses represent standard deviations.
Scores range from 20 to 60. Maximum score = 32 for each cell.
Figure Captions

Figure 1. Sequence of events for the experimental task.
Low Memory Load Condition

1.5 sec

1.5 sec READY?

6.0 sec

Memory Load Set
A B C D E F

Memory Load Set
E A C B F D

Mental Arithmetic Task
53 + 69 = ?

Mental Arithmetic Task
80 – 27 = ?

Answer to Mental Arithmetic Problem

Answer to Mental Arithmetic Problem

Memory Test Set
A B C D E F

Memory Test Set
E A C B F D

OR

OR

Recording of Memory Load Task RT

Recording of Mental Arithmetic Task RT

F E D C B A

E D C B F A