Doctoral Dissertation

An Ecological Validation of the Child Attention Network Test (CANT) in Korea

Department of Education

GRADUATE SCHOOL
JEJU NATIONAL UNIVERSITY

Jeong, Gyeongmi

February, 2021

An Ecological Validation of the Child Attention Network Test (CANT) in Korea

Jeong, Gyeongmi

(Supervised by Professor Song, Jaehong)

Submitted to the Graduate School of Jeju National University in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Education

December, 2020

This dissertation has been examined and approved by Doctoral Committee:

Committee Chair ____________________

Committee Member ____________________

Committee Member ____________________

Committee Member ____________________

Committee Member ____________________

December, 2020

Department of Education

GRADUATE SCHOOL

JEJU NATIONAL UNIVERSITY

Abstract

An Ecological Validation of the Child Attention Network Test (CANT) in Korea

Jeong, Gyeongmi

Department of Education, Jeju National University, Korea

Supervised by Professor Song, Jaehong

In order to identify children's attention problems and develop intervention programs, it is necessary to use a valid and reliable attention assessment. This research attempts to examine the ecological validity of the Child Attention Network Test (CANT), which was developed based on the attention network theory and has been widely used in various academic fields. By doing so, this study aims to build a ground for further attention research and gain insights of how attention operates in academic activities. To achieve the goals, Study 1 tests the item quality, reliability and validity of the CANT with Korean samples in order to find out whether the psychometric properties of the CANT would be replicated in samples with a different cultural background. Study 2 examines the relationship between the performances in the CANT and the mathematical word problem-solving in order to demonstrate the generalizability of the CANT performance to that of mathematical word problem-solving.

In Study Ⅰ, 71 elementary students in grades one to six were sampled in a province of Korea to take the CANT. Thirty-five students in grades four to six

among the sampled population took another performance-based attention test, FACT-Ⅱ-SR, and took the CANT 10 to 14 days later for the second time. Their parents and homeroom teachers were asked to rate their attention-related behaviors on KADHDDS or K-ADHDRS-T. Analysis of item discrimination and validity were conducted to test the quality of the items. Reliability was examined by calculating Cronbach's alpha, the split-half reliability test, and the test-retest reliability test. Validity was tested in terms of its concurrent validity. As a result, the items of the CANT appeared valid and contributing to the test's discrimination. The reliability of attention network scores ranged from insufficient to excellent. Some of the CANT scores showed significant correlations with the scores acquired from the criterion measures.

In Study 2, 31 students in grades five to six in a province of Korea were sampled. They took the CANT and solved mathematical word problems. In order to see the CANT's predictive power on the math performance, Pearson correlation analysis, hierarchical regression analysis, and visualizing the performance patterns were performed. It was found that the accuracy on the CANT predicted the accuracy on the mathematical word problem-solving. It also predicted the speed of mathematical word problem-solving mediated by the computation skill. Although alerting, orienting, executive control networks did not appeared to significantly predict the performances on the word problems with irrelevant information and inconsistent relational terms, the attention network scores showed significant correlation with them. Additionally, the accuracy and speed pattern of the CANT appeared similar to that in the mathematical word problem-solving. These results imply that the CANT is an ecologically valid measure, thus increasing its usefulness in school settings.

This research contributes to building ground for further attention research by testing the CANT's psychometric properties in Korea. The results indicate that the CANT consists of valid and discriminable items and is a valid measure. Some of the network scores are reliable, whereas other scores are required some improvement in their consistency. The results also demonstrate that the CANT performance represents the performance in mathematical word-problem solving. Therefore, it could be concluded that the CANT is an ecologically valid attention measure, which can be employed in Korea and educational settings. Further research would continue to validate the CANT and come up with ideas to improve its reliability. Moreover, based on the insight, how the attention networks impact students' school adjustment should be examined in more depth and in more various school fields in the future.

Table of Contents

List of Tables

List of Figures

Ⅰ. Introduction

1. Research Background

Since attention plays a crucial in the learning process, many researchers and clinicians have developed and validated various attention assessments for students in Korea. Especially, as inattentiveness rather than hyperactivity and impulsivity appeared to significantly predict the lower academic achievement of ADHD students than their normally developing peers (Jaekel, Wolke, & Bartmann, 2013), a sound attention measure that can diagnose inattentive symptoms and evaluate the effects of intervention programs has been demanded in the academic field and educational settings. As a result, researchers and clinicians in Korea have validated several attention assessments developed in other countries. For example, Oh and Lee (1989) validated the Conners Rating Scale-Short Version in Korea, and Choi, Kim, and Park (1994) validated the Test of Variables of Attention (T.O.V.A) among Korean samples. Since these attempts, other attention measures such as Attention Deficit Disorder Evaluation Scale-School Version and ADHD Rating Scale have been validated in Korea (Bae et al., 1997; So et al., 2002). Furthermore, other researchers, such as Kim and Kim (2008), have developed new attention assessments for students. These assessments have been utilized in studies to discriminate students with attention problems and test the effects of attention intervention programs (Jeong & Song, 2019).

However, findings from a meta-analytic review on psychometric properties of attention measures in Korea indicated that there should be caution in using attention measures. According to Jeong and Song (2019), the test-retest reliability of performance-based attention measures such as T.O.V.A reached the level of clinically acceptable, and the effect size of concurrent validity was small. Considering the frequent uses of performance-based tests such as the Continuous Performance Test (CPT) in academic fields and clinical settings, the result leads to concern. They also mentioned that they had failed to find enough studies on the development of attention assessments reporting multiple reliability and validity indices. Therefore, it would be a valuable work to find another performance-based attention measure and test its validity and reliability among Korean students.

Recently, attention has been studied from the neurocognitive perspective (Jeong, Song, & Kim, 2019). In the early years of attention literature, experimental psychological tests were used in attention research since Welford (1952) had found the bottleneck phenomena in the process for two consecutive signals resulting in a response time gap (Styles, 2006). Later, studies on patients with brain damage and the development of brain imaging methods took attention research to the next step. Many researchers have been utilizing neuropsychological tests and brain imaging technologies to uncover particular brain structures and pathways related to attention. This transition in attention literature led researchers to develop attention measures based on the neuropsychological results (Fan et al., 2002; Mirsky, Pascualvaca, Duncan, & French, 1999; Posner & Petersen, 1990).

Among the attention measures taking the neuropsychological approach, the Attention Network Test (ANT) has gained attention in the literature. It was developed based on the attention network theory suggested by Posner and Petersen (1990). According to the theory, there are anatomically independent structures of attention networks from other cognitive processing systems. The network consists of structures of alerting, orienting, and executive control. They are independent from each other but operate interactively. Later, Fan et al. (2002) developed the ANT by modifying a flanker task and a signal detection task, which had been found to activate the attention networks. Rueda et al. (2004) revised the ANT into a more child-friendly, and named it as the Child Attention Network Test (CANT). The tests have been used in many neuropsychological studies and genetic studies, thus expanding the attention literature (MacLeod et al., 2010).

The ANT and the CANT have been used in studies in Korea. Kim (2019) assessed the orienting attention of people with sluggish cognitive tempo through the ANT. Song (2007) analyzed the instability of the attention networks of schizophrenic patients. Kim & Youn (2013) used the CANT investigating the difference in preschooler's attention according to their gender and cognitive style. Even though the ANT has been contributed to many academic fields, it has not been employed in Korea as frequently as in other countries. Furthermore, the CANT has not been often used in studies, thus leading to further investigation for its utility.

One of the reasons for the infrequent usage of the CANT is that its psychometric properties have not been thoroughly investigated. In order to assure the soundness of a measure, multiple psychometric properties have to be continuously examined. Validity shows how well a test measures what it is supposed to measure (Cohen, Swerdlik, & Sturman, 2013). The evidence of the validity of a test can be found through construct validity analysis, concurrent validity analysis, and factorial validity analysis. The CANT's validity has been tested in terms of its test design. According to Rueda et al. (2004), the effects of the cue and flanker manipulation were valid, resulting in different response times and accuracy. However, no studies have investigated the CANT's other validity, such as concurrent validity and predictive validity. Reliability refers to the consistency of the test result. Internal consistency and stability of the result are some of the reliability evidence. Rueda et al. (2004) reported the split-half reliability of the CANT ranging from .02 to .94 and found out no significant correlations between the original scores and the scores measured 6.5 months later. Individual item should be assessed in terms of its contribution to the test. The items of the CANT, however, have not been examined in previous studies. On the other hand, the ANT's psychometric properties have been tested in many ways. Its validity has been investigated in neuroimaging studies as well, supporting the existence of three attention networks (Fan et al., 2005). The split-half reliability of the ANT was tested in multiple studies (MacLeod et al., 2010), and there have been attempts to improve the ANT's test-retest reliability (Wang, Fuchs, & Fuchs, 2016). In Korea, only Lee et al. (2017) examined the ANT's test-retest reliability with schizophrenic patients. No studies attempted to test the psychometric properties of the CANT in Korea, thus requiring further investigation.

Meanwhile, there has been an increasing demand for ecological validity evidence for neuropsychological tests (Chaytor & Schmitter-Edgecombe, 2003). Ecological validity refers to the extent to generalize the results acquired in a condition to other conditions (Korean Educational Psychology Association, 2014). If the outcome of a test can not be reproduced in different contexts and transferred to other settings, the test would be determined ecologically invalid.

Chaytor and Schmitter-Edgecombe (2003) have identified two methods to evaluate the ecological validity of a test: verisimilitude and veridicality. Verisimilitude shows how much a test's cognitive demands resemble those in real-life environments. Cognitive tests involving experimental tasks that simulate real-life tasks can be considered ecologically valid in this respect. The Rivermead Memory Test which assesses memory by requiring test takers to associate names with faces and remembering appointments is an example of a verisimilarly valid test. However, since this approach aims to identify people struggling in everyday life regardless of the reasons for the struggle, it might fail to detect the precise sources of real-life problems and identify people performing properly in everyday life but still neurocognitively impaired. On the other hand, veridicality refers to the extent to which everyday functioning is associated with the result of a test conducted in a controlled environment (Ready, Stierman, & Paulsen, 2001; Wallisch, Little, Dean, & Dunn, 2018). Neuropsychological tests often happen in laboratory settings with artificial stimulus sampling only a small portion of behaviors from the test takers. Therefore, it is vital to demonstrate the representativeness of the test situations and stimulus and the generalizability of the test performance to real-world functioning. Veridicality is tested by examining the relationships of test performances and real-world performances such as employment status (Chaytor and Schmitter-Edgecombe, 2003). Since these ecological questions are of great interest to examinees and examiners, neuropsychological tests are demanded to be evaluated their ecological validity (Chaytor and Schmitter-Edgecombe, 2003; Kibby, Schmitter-Edgecombe, & Long, 1998).

The reproducibility of the test result across different cultures is another aspect of the ecological validity of a test. When researchers try to utilize a test that was developed within a culture, its psychometric properties in the target population with different cultural backgrounds should be ensured. It is because individuals' experience highly depends on the culture they belong to, and psychological constructs are likely to be influenced by the culture (Mushquash & Bova, 2007). Cognitive tests could yield different results among populations with different cultural backgrounds due to the differential familiarity with testing methods (van de Vijver & Tanzer, 1997). For example, technical equipments other than paper-and-pencil could result in cross-cultural differences due to unexpected variables such as curiosity. In addition, test stimulus and response procedures could cause confusion in some cultures (van de Vijver & Tanzer, 1997). Because of these culture-sensitive factors, test results can be biased.

Furthermore, some cognitive tests favor test takers from a certain cultural group. Aboriginal children showed lower performance on the Wechsler Intelligence Scale for Children (Mushquash & Bova, 2007). Chinese children showed a better performance in visuospatial processing than Greek children due to Chinese children's extensive visuospatial practice in learning to write their mother tongue (Demetriou et al., 2004). Valencia, Suzuki, and Salinas (2002) examined 62 studies investigating IQ of White and minority ethnic groups and found out that 30% of the studies showed culturally biased psychometric properties of intelligent tests, although 70% of the examined studies yielded non-biased findings. Therefore it is necessary to conduct a psychometric analysis before using a test for a cultural group that is different from the group the test was initially intended for.

The ecological validity of the CANT and the ANT have not been under investigation. In order to maximize the utility of the CANT, its ecological validity needs to be examined in a variety of settings. Since it was devised in another country for children with a different cultural background, it is necessary to see if the test yields the expected outcome in Korea as well. Although the CANT does not include any linguistic information during the experimental blocks, the manipulation of stimulus, cues, and flankers requires to be validated in Korean settings. In addition, reliability should be assessed before its use with different populations because the familiarity of test formats and instructions can be sources of test errors (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014).

The veridical validity of the CANT also needs to be tested. As CANT taps the attention networks during the test, the performance of it should be able to make inferences on cognitive performances involved in academic achievement. Attention operates in the process of encoding and retrieving information. It also has an influence on saving and manipulating information in the working memory (Long, Kuhl, & Chun, 2018). While learning, students need to allocate the right amount of attention resource to the most important stimulus and reduce the amount of attention for completing a task by automatizing necessary skills (Brunning, Schraw, & Norby, 2011). Therefore, students with attention deficits are expected to demonstrate difficulties in their learning process.

Especially, Mayer (1980) has described how attention influences mathematical problem-solving. He suggested that problem solvers go through four phases during the mathematical word problem-solving, of which each phase requires problem solvers to choose only relevant information while suppressing irrelevant information. Although he did not provide quantitative evidence for attention operation during mathematical problem-solving, other researchers have been attempting to demonstrate the relationship between attention and mathematical achievement (Anobile, Stievano, & Burr, 2013; Ansari, Lyons, van Eimeren, & Xu, 2007). It indicates that attention measures for children should be able to predict mathematical problem-solving to be considered ecologically sound.

Neuropsychological tests were developed to detect brain lesions and identify specific brain structures and their unique functions during cognitive tasks (Chaytor and Schmitter-Edgecombe, 2003). The CANT and the ANT have been proved to effectively activate the distinctive brain structures related to attention networks (Fan & Posner, 2004). Therefore, it is natural to ask if the performance in this laboratory task with the artificial stimulus in a controlled experimental setting would represent the performance in the real world, in this study, in the educational situation. A meaningful relationship between the performances in the two different settings would demonstrate the CANT's representative nature to the real world.

2. Purpose and Research Questions

This study aims to test the ecological validity of the CANT in Korean elementary students. Since ANT has gone through several psychometric tests in multiple cultures, it will be meaningful to investigate whether the CANT shows consistent and valid psychometric features in Korean samples. Furthermore, its predictable power on mathematical problem-solving would increase its ecological validity demonstrating its utility in educational settings. By achieving the goals, this study would contribute to the academic field and educational settings by providing a valid and reliable attention measure and to expanding attention research to developmental psychology and educational psychology. The research questions are as follows:

Research Question 1: How valid is the CANT psychometrically in the setting of Korean elementary school?
Research Question 2: Does the CANT account for individual differences in mathematical word problem-solving?

3. Definitions of Terms

1) Attention

Attention is a cognitive resource needed in selectively perceiving and consciously processing information. It consists of obtaining and maintaining the alerted state of mind, orienting specific stimuli relevant to the task goals, and resolving cognitive conflict.

2) Attention Network

Attention network refers to the network between anatomically independent localized brain structures that operate interactively in attention-needed tasks. There are three attention networks, which are alerting, orienting, executive control networks. Alerting is to create and maintain an optimally alerted mental state in order to perceive and process stimuli or information. Orienting is to select and locate specific stimuli and shift attention to them, excluding distracters. Executive controlling is to consciously focus attention on the goal-driven stimuli resolving conflicts.

3) Ecological Validity

Ecological validity refers to the extent to which a test score can predict the performances in different settings, such as in different cultures or non-experimental situations.

4) Item and Item Quality

Item refers to the combinations that resulted from the cue types and flanker types in the CANT. Since there are four cue types and three flankers types in the CANT, 12 different items are yielded in this study. Item quality refers to the validity of each item in relation to the test. In this study, item quality is defined in terms of two aspects of each item. First, it is defined as the degree of correlation of the response time and accuracy of each item with those of the test. Second, it is defined as the degree of cue and flanker manipulation effects on the response time and accuracy of the individual item.

5) Validity

Validity refers to the degree of how well a test measures what it is supposed to measure (McCowan & McCowan, 1999). In this study, validity is defined as the degree of correlation between the scores of CANT and the criterion tests.

6) Reliability

Reliability refers to the degree of consistency and stability of a test score. In this study, it is defined as the internal consistency assessed by Cronbach's alpha and as the correlation between the scores of CANT across three test blocks. It is also defined as the correlation between the test scores measured at two different times.

Ⅱ. Literature Review

1. Attention

Although attention is used by anyone but hard to define (Lee & Park, 1998; Styles, 2005). William James, the founder of modern psychology, stated:

> "Everyone knows what attention is. It is the taking possession by the mind in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thoughts. Focalization, concentration of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others, and is a condition which has a real opposite in the confused, dazed, scatterbrained state which in French is called distraction, and Zerstreutheit in German".
>
> (James, 1890: pp.382-383)

Attention was not the main topic for researchers in the behaviorism era due to its invisibility. However, it emerged as the main research topic in the 1950s, during which information processing theory began to dominate the academic field of psychology. Since then, many researchers have tried to define attention as a part of information processing.

The development of information processing theory, along with rigorous psychological experiments, formed the current definition of attention. Broadbent's filter theory of attention is based on the single-channel approach to information processing, explaining that attention plays a role as a filter or a bottleneck somewhere in the information processing chain to prevent from overloading of stimuli to be processed (Broadbent, 1958: as cited in Hancock, Oron-Gilad, & Szalma, 2007). Whereas Broadbent focused on the structural limitation in the process, Kahneman (1973) suggested attention in terms of cognitive capacity. He suggested that a certain degree of processing capacity could be allocated to a task, thus perceiving processing capacity as pools of energy. Humans need to allocate a certain amount of mental resources, i.e., attention, to a target stimulus due to the limited amount of cognitive capacity they possess. In this context, Brunning et al. (2011) defined attention as deploying cognitive resources to a currently working task appropriately. Anderson (2011) defined attention as a potential process of allocating cognitive resources in need of competitive information processing. Das (2002) suggested that attention was a mental process by which a person registers important stimuli and resists irrelevant information. Sheridan (2007) defined attention as the focusing of sensory, motor, and mental resources on aspects of the environment to gain knowledge. However, he admitted that he could not explain what "focus" really means in the nerve system. These definitions see attention as mental fuel for guiding the information process to a target outcome (Bruning et al., 2011; Lee & Park, 1998). Although there are still ongoing arguments of whether it is a single resource to be allocated or multiple resources distributable to different aspects of stimuli at once (Hancock et al, 2007), the conceptions of attention as a kind of energy or as a structural filter have greatly contributed to defining attention.

Attention is also considered as a voluntary control exerted over a more automatic brain system (Posner & Petersen, 1990). In this view, attention is regulating multiple brain networks by attentional networks associated with maintaining the alert state, orienting, or resolving conflict (Posner & Rothbart, 2007). Some behaviors might not accompany attention, therefore called automatic actions, while novel actions required to prevent habitual actions demand deliberate attentional intervention. This approach focuses attention during the process of planning, initiating, and executing actions (Norman & Shallice, 1980). Allport (1987) and Neumann (1987) conceived attention as an agency for selecting for action (cited from Wu, 2014). Attention in the control of action driven by goals and instructions is a growing research area in attention literature (Styles, 2006).

Meanwhile, Wu (2014) conceptualized attention from multiple perspectives. His first approach is the function-centered approach. On this approach, he defined attention based on what attention is for. That is, attention functions as a filter for information for further processing, binding features in representing objects, a spotlight highlighting its target, and selecting targets for memory, consciousness, or action (Wu, 2014). The second approach is a mechanism-centered approach in which the emphasis is on the nature of the process of selection. In this approach, attention is something that emerges from competition for limited resources and in the process of preparing a motor response. At a more concrete neural level, attention is the modification of neural signals, and alerting the brain area to which a neuron responds (Wu, 2014). The third approach is the phenomenology-centered approach. On this approach, attention is considered as a gatekeeper for selection for action rather than as a distinctive way of consciousness. He claims that attention is conscious to the extent, although there is much work to be done to support the notion (Wu, 2014). In conclusion, attention can be defined as a mental operation in information processing, which includes selectively perceiving, consciously processing, and taking action to stimuli. It is necessitated by a limited processing capacity.

Although researchers defined attention differently, they have reached an agreement that attention is not a unitary construct but an umbrella term indicating multiple aspects of cognitive operations (Mirsky et al., 1999; Styles, 2006). Kim and Kim (1999) summarized six attentional components, which are alertness, neglect, focused attention, divided attention, sustained attention (vigilance), and executive control of attention. Alertness is the most basic condition to maintain attention. It is a condition of the central nervous system, which affects general reactions to stimuli. Neglect is a phenomenon caused by a disruptive function of spatial attention resulting in failure to perceive and respond to stimuli in the contralesional field (Gabrieli, Whitfield-Gabrieli, 2007). They first introduced selective attention and divided it into two categories: focused attention and divided attention. Selective attention is the ability to pay attention intentionally to a stimulus among multiple stimuli presented at the same time or a certain aspect of a stimulus. Focused attention determines the vector of attention, whereas divided attention determines the amount of attentional capacity allocated. Sustained attention is the capability to maintain the state of attentiveness for a certain period of time. Executive control of attention is the ability to inhibit the response to irrelevant stimuli, decide when to shift attention to a different environmental aspect, and determine the order of responses.

Neuropsychological approaches tried to reveal the mechanism of attention in the brain, thus suggesting procedural components of attention when it is operating. For example, Pribram & McGuinness (1975) proposed three separate but interacting neural systems of attention: arousal, activation, and effort. Arousal is defined as a phasic physiological response to stimuli. The activation is defined as a tonic physiological readiness to respond. Lastly, an effort is demanded to coordinate arousal and activation. Mirsky et al. (1999) considered attention as a set of processes supported by different brain regions that compose a particular system for multiple attentional functions: focus/execute, sustain/stabilize, shift, and encode based on the result of a factor analysis of neuropsychological test data. Focus/execute refers to the ability to focus attentional resources inhibiting non-relevant stimuli. Sustaining refers to the capacity to stay on task vigilantly for an appreciable time. Attention shift refers to shifting attentional focus from one aspect of a stimulus to another flexibly and efficiently. The last component is to encode, which refers to a mnemonic capacity to keep information in mind for a brief time while performing some action or cognitive operation on it.

Knudsen (2007) suggested four fundamental process components to attention: working memory, top-down sensitivity control, competitive selection, and bottom-up filtering for salient stimuli. According to him, filtering for salient stimuli refers to selecting information based on their distinctive properties that are likely to be important. This is often called bottom-up attention. Competitive selection occurs to select information for later representation and entry to working memory. In this process, the strength of signals is compared, and the effects of distracting stimuli are eliminated. Working memory stores a limited amount of information for a short time. Evaluating and manipulating the stored information take place for decision making and planning for actions. The information of the attended stimuli enters working memory and is represented and processed. Top-down sensitivity control occurs to improve the quality of the information that is being processed during competitive selection and working memory. The mechanism for this process involves eye movements and other orienting behaviors. As outlined, attention is an overarching term indicating multiple cognitive elements. The components of attention suggested by researchers are summarized in Table 1.

Table 1

Components of Attention

Components	Alertness	Selective Attention	Sustained Attention	Divided Attention	Attention Shift	Attentional Control	Others
Pribram & McGuinness (1975)	● (arousal)		● (activation)			● (effort)	
Kim & Kim (1999)	●	● (focused attention)	●	●	●	● (executive control)	● (neglect)
Mirsky et al (1990)		● (focus/execute)	● (stabilizing)		●		● (encode)
Das (2002)		● (selection & focusing)			●	● (resisting distraction)	
Knudsen (2007)		● (competitive selection)				● (top-down sensitivity control)	● (working memory, automatic bottom-up filtering)
Atkinson, Braddick, & Breckenridge (2010)		●	●			●	
Mahon & Schneider (2012)		●	●	●		●	
Petersen & Posner (2012)	●	● (orienting)				●	

2. Attention Network Theory

One of the most prominent theories of attention is the attention network theory suggested by Posner and Petersen (1990). The theory was inspired by Hebb's network view (Petersen & Posner, 2012; Posner & Petersen, 1990; Posner & Rothbart, 2007). Hebb and Hebb (1949) suggested the cell assembly theory and the concept of phase sequence, which is involved in the coordination of multiple cell assemblies underlying human thoughts, feelings, and actions. He further argued that activity flow in a set of interconnected neurons represented psychological event, sensation, expectation, emotion, or thought (Posner & Rothbart, 2007). Taking this network approach, Posner and Petersen (1990) suggested the attention network theory to explain attention, the cognitive psychological phenomenon, in the integration with neuroscience.

Empirical evidence contributed to the development of the theory. Behavioral studies of normal adults and patients with varying forms of brain injuries supported the theory. For example, initial imaging studies that observed the brain activity with positron emission tomography showed, patients with lesions centered in the superior temporal gyrus could regulate attention allocation to global and local levels, while patients with lesions in the rostral inferior parietal lobule could not (Robertson, Lamb, & Kight, 1988). Research findings from galvanic skin responses in humans and monkeys, heart rate responses to warning signals, cerebral blood flow, and metabolism involving vigilance tasks provided a uniform implication about the role of the right hemisphere in attention (Posner & Petersen, 1990). Such accumulated empirical evidence has been devoted to the development of the theory.

New methodologies advanced the theory. Neuroimaging technologies provided the discrete anatomical basis of the attention system. The mapping of the human genome contributed to an increased understanding of individual differences in the efficiency of the attentional system (Posner & Rothbart, 2007). The development of research methods is expanding attention research in the field of cognitive psychology and the field of developmental psychology, physiology, and genetics.

According to the attention network theory, attention is regarded as an organ system with its own anatomy, circuitry, and set of functions (Fan & Posner, 2004). Attention network is anatomically separate from other systems for information processing, such as encoding incoming stimuli, making decisions, and producing outputs, although they still interact. There are three networks within the attentional network, which are anatomically independent brain areas from each other. They carry out a different set of attentional processes, though operate in an interactive manner in tasks requiring attention.

Attention network can be subdivided into three networks: alerting, orienting, and executive control networks (Fan et al., 2009; Petersen & Posner, 2012; Posner & Rothbart, 2007). Alerting is achieving and maintaining a state of vigilance to prepare for later processing and to optimize performances. Alerting can be divided into phasic alertness and tonic alertness. Phasic alertness is the ability to increase response readiness to an incoming stimulus following a warning signal. Tonic alertness is the ability to stay wakeful and aroused during rather a long boring task. The alerting network is lateralized to the right hemisphere, associated with thalamic, frontal, and parietal regions of the cerebral cortex. Norepinephrine (NE) is known to influence the operation of the network.

Orienting refers to the selection of information from sensory inputs. It includes a rapid or slow attention shifting among objects within a sensory modality or among multiple modalities. Attention shifting consists of three basic operations: disengaging attention from its current focus, moving attention to a new target, and engaging attention at the new target (Berger & Posner, 2000; Fan et al., 2009). When orienting is accompanied by head or eye movements toward the target, it is overt orienting. On the other hand, if it does not involve any movement of the head or eyes, it is referred to as covert orienting. Sometimes a sudden target event can direct attention to its location (exogenous orienting or reflexive orienting), while a more voluntary orienting can occur locating for a target (endogenous orienting). Orienting is associated with specific brain structures such as the superior parietal lobe, temporal parietal junction, and the frontal eye fields. It is modulated by the neurotransmitter acetylcholine (Posner & Rothbart, 2007).

The executive control network deals with a more complex operation. It is involved in monitoring and resolving conflicting thoughts, feelings, and responses. It is related to the control of goal-directed behaviors, target detection, error detection, inhibition of automatic responses, overcoming habitual responses, and executing a novel or not well-learned response. Thus, executive attention moderates the activity of sensory, cognitive, and emotional systems (Berger & Posner, 2000; Peterson & Posner, 2012). Brain areas of anterior cingulate, lateral ventral, prefrontal and basal ganglia are associated with the executive attention operation. The neurotransmitter dopamine modulates the operation of it (Posner & Rothbart, 2007).

The three attentional sub-networks are structurally distinct networks with different functions, however, they interact among themselves. To process information deeply in real-world situations, manipulating and maintaining the alerted state, orienting sensory or mental focus to the target information, and top-down attentional control should operate simultaneously. Correlations between the effects of three networks have been detected in experiments (Callejas, Lupianez, Funes, & Tudela, 2005; Fan et al., 2009). The research findings suggested that depending on the condition, alerting can exert a negative influence on executive attention and facilitate the orienting network.

Orienting facilitates executive attention (Gallejas et al., 2005; Ishigami & Klein, 2009). However, there is still much to learn about the interplay among them.

There are individual differences in the ability to attend. Differences in the efficiency of alerting, orienting, and executive attention networks were detected among individuals (Fan et al., 2002; Ishigami & Klein, 2009; Rueda et al., 2004). Neuroimaging evidence showed significantly different activation in the brain areas associated with the networks between people. Variations in genes related to the neuromodulators such as dopamine for executive attention appeared to contribute to the risk of developing attention-deficit related disorders (Fossella et al., 2002). Considering that the dopamine 4 receptor gene interacts with parenting styles, environmental factors also play a role in attentional development. This implies that intervention programs could improve the efficiency of attentional networks since the genetic and environmental factors continue to influence throughout adulthood (Larsen et al., 2010).

3. Measuring Attention

As attention can not be simply defined, it is not easy to assess. Furthermore, a single event in real-life situations mostly includes several aspects of attention operating interactively and simultaneously. Therefore, many researchers have tried to test attention as a unitary construct, while others have attempted to find ways to manipulate tasks tapping each component of attention.

Assessment of attention can be divided into three categories: behavior rating scales, performance-based tests, and structural interviews (Jeong & Song, 2019; Mahone & Schneider, 2012). Rating scales require participants to quantify the frequency or degree of their attention-related behaviors in various life situations. If it is a child whose attention is being assessed, parents and teachers are asked to rate his or her behaviors. On the other hand, performance-based measures require participants to respond to attention-needed tasks, yielding performance data on accuracy and response time. Structural interviews are conducted by licensed clinicians asking structured questions and observing interviewees' behaviors. It is encouraged to use all three types of attention measures to obtain a comprehensive understanding of one's attention for clinical uses.

Performance-based attention measures have some advantages over the other types of assessment. They minimize the influence of raters' bias on the performance by applying tests directly to participants. It is also possible to employ neuroimaging methods during attention tasks so that researchers can get a clear picture of the blood flow or electrical activation occurring in the brain. Also, they are easy to administer in various settings (Nichols & Waschbusch, 2004). Especially, computerized tasks yield a millisecond-unit respond time result enabling testers to measure constructs that are detectable only at a subtle level. These benefits led researchers to developing performance-based attention tasks.

The most frequently used performance-based attention measure is the continuous performance test (CPT) (Mirsky et al, 1999). It was initially developed by Haldor Rosvold, one of the pioneers in attention literature, who served as a psychologist in the Canadian Army in World War Ⅱ. He noticed attentional problems occurring to the frontal lobes-injured soldiers and devised the CPT to assess their attention (Mirsky et al., 1999). The CPT requires subjects to watch a visual display for a designated time and press a response key whenever a target appears. The original form of the CPT has been modified by researchers later on. For example, in the original version, subjects were given an X as a target to respond. In contrast, in the later versions, subjects are required to respond to a non-literate stimulus, a degraded version of X, an X following an A, or auditory stimuli. The CPT measures have been used by researchers as well as clinicians (Mirsky et al., 1999; Nichols & Waschbusch, 2004)

Not only the CPT, but there are also other assessment paradigms of attention. Visual cancellation tasks require subjects to locate and mark the targets among a visual array as quickly as possible. Digit span measures require subjects to recall series of numbers forwards or backward. In Delay of gratification tasks, experimenters present subjects, usually children, with an attractive toy or snack and instruct them to wait until the experimenter permits them to either play with or eat the target, thus measuring inhibitory control (Mahon & Schneider, 2012; Nichols & Waschbusch, 2004).

The Attention Network Test (ANT) was developed to test the efficiency of the attentional network. A cued detection task and a flanker task were combined with some manipulation on cue presenting and flanking. The efficiency of the alerting network is assessed through a cued detection task accompanying warning cues. It assumes that the warning cues prior to a target enhance the response speed. Orienting is measured by manipulating a cue indicating where the target is likely to occur, thus directing attention to the cued location. In order to assess executive attention, a flanker task was employed in which incongruent flankers, that is, conflicting distracters, sometimes appear around the target requiring participants to resolve the conflict (Fan et al., 2002; Petersen & Posner, 2012; Posner & Rothbart, 2007).

Psychometric properties of attention measures such as reliability and validity have been evaluated in many ways. Reliability is the consistency of the assessment scores among items, times, and examiners (Cohen et al., 2013). Rating scales assessing attention-related behaviors have been estimated using consistency measures such as Cronbach's alpha, split-half reliability, test-retest reliability, and inter-rater reliability. Cronbach's alpha of rating scales that were developed or validated in Korea appeared to be excellent. Test-retest reliability of performance-based attention measures such as CPT and FACT ranged between .17 and .92 (Jeong & Song, 2019). This result is in line with CPT measures' reliability ranging between .14 and .94 (Ogundel, Ayyash, & Banerjee, 2011).

Validity shows how well a test measures what it is supposed to measure (Cohen et al, 2013). Assessments of attention have been tested on their validity in terms of content validity, concurrent validity, and construct validity. Concurrent validity of rating scales of attention in Korea appeared to show a high correlation with other criterion tests, whereas performance-based attention measures showed a low correlation with criterion tests (Jeong & Song, 2019). This result is consistent with the result of studies where various laboratory attention measures were reviewed in terms of their validity (Nichols & Waschbusch, 2004). The CPT appeared to show a moderate correlation ranged between .23 and .41 with criterion measures. It appeared to effectively discriminate ADHD children from normal equivalents. Although there is no ground rule for the minimum level of concurrent validity estimates, the low to moderate level of concurrent validity of performance-based attention measures led researchers to advise to use them along with other attention measures with caution (Nichols & Waschbusch, 2004).

The reliability of the ANT and CANT have been tested in earlier studies. Split-half reliability of the ANT was initially reported as .52 for alerting, .61 for orienting, .77 for executive control (Fan et al., 2002). MacLeod et al. (2010) collected split-half reliability data of 1,141 participants in 15 unique studies of healthy individuals and yielded reliability estimates of .65 for executive control, .32 for orienting, and .20 for alerting. The split-half reliability of CANT in Rueda et al. (2004)'s study was .59 for executive control, .02 for orienting, and .37 for alerting, whereas it was .94 for overall response time and .93 for error rate. Test-retest reliability of alerting, orienting, and executive control scores in the ANT ranged between .35 and .80 (Wang et al., 2015). However, Rueda et al. (2004) found no significant correlations between the original scores and their repetition which were measured 6.5 months later. Reliability of the ANT has been reported to be high for orienting and executive control and moderate for alerting in the study with schizophrenic patients and moderate for all the three networks for healthy controls in Korea (Lee et al., 2017). The CANT has been used in a study with samples of young children aged 4 to 5 in Korea, but any reliability estimates have not been reported, thus demanding a further attempt to test its reliability (Kim & Youn, 2013).

The validity of the ANT has been well-outlined in previous studies, whereas the CANT's validity can be inferred from the validity estimates of the ANT. Since the flanker test and the cued detection task, which are combined in the ANT, are well-established psychometric measures, the face validity has been considered as satisfactory (MacLeod et al., 2010). Neuroimaging studies have supported the presence of three attention networks activating different brain regions (Fan et al., 2005; Posner & Rothbart, 2007). The effects of cue and flanker manipulation corresponded with what they had been expected (Fan et al., 2002). A significant main effect of cue type and flanker type on the response time was found in the CANT (Rueda et al., 2004). Other than this, however, the validity of the CANT has not been tested explicitly.

The ecological validity of the CANT has not been examined in previous studies. Although CANT has been used in studies with samples from Korea, Ireland, and Norway, none reported any psychometric properties of the tool. Hence it is not clear whether the tool shows similar psychometric properties in different cultures other than the U.S. in which the CANT was initially developed and validated in Rueda et al. (2004)'s study. In addition, no significant differences in the efficiency of attentional networks have been found between ADHD children and typically developing children (Adólfsdóttir, Sørensen, & Lundervold, 2008), which indicates the inability of the CANT to predict inattentive behaviors in daily lives. Therefore, a further examination of the ecological validity of the CANT is required to ensure its utility in diverse settings.

4. Attention and Mathematical Achievement

Attention-related skills are essential to school success (Martin, Razza, & Brooks-Gunn, 2012). As students have a limited cognitive resource available at a time, they have to carefully choose stimuli to process further, while inhibiting irrelevant stimuli such as noise in the hallway, whispers from behind, and falling leaves out of the window. Besides, in order to solve problems requiring critical thinking skills or applying already-learned knowledge, students need to minimize attentional resources allocated on simple skills and factual knowledge. For example, if students want to read something for meaning, they have to spend little or no attention to word decoding. By automatizing basic skills, students can devote attention to more challenging learning tasks. Thus, skilled learners selectively attend to important information and automatize skills to conserve resources (Bruning et al., 2011).

Attention is involved in multiple aspects of mathematical processing. Using the framework of information processing model, Mayer (1980) suggested that attention plays a role when a problem solver needs to choose important information, and suppressing irrelevant information while word problem-solving. Sustained visual attention measured by an object tracking test predicted mathematical achievement assessed by tasks such as Arabic numeral reading, simple calculation, counting, and complex written calculation (Anobile et al., 2013). Number processing involves attention, as well. Functional neuroimaging studies showed that number processing depended on different attentional networks; small numerosities engage the brain region related to stimulus-driven attentional network, and large numerosities engage with the brain area related to goal-directed attentional network, suppressing the other at the same time (Ansari et al., 2007). Therefore, children with attention problems are expected to demonstrate lower performance in mathematics compared to typically developing peers (Hart et al., 2010; Martin et al., 2012; Zentall, Smith, Lee, & Wieczorek, 1994).

Research has demonstrated correlations between attention and achievement of math word problem-solving. Studies employing attention-related behavior rating scales by parents or teachers have reported the significant negative correlation between inattentive behaviors and performance of mathematical word problems (Fuchs et al., 2006; Rogers et al., 2011). Sustained attention assessed by a visual tracking task appeared to predict word problem-solving performance (Anobile et al., 2013). Although not many studies have been accumulated, recent studies employing neuropsychological measures such as functional magnetic resonance imaging technique and laboratory attention measures such as continuous performance test and the ANT have suggested plausible evidence of correlation between the two latent variables (Anobile et al., 2013; Ashkenazi & Silverman, 2017; Mathieu et al., 2018).

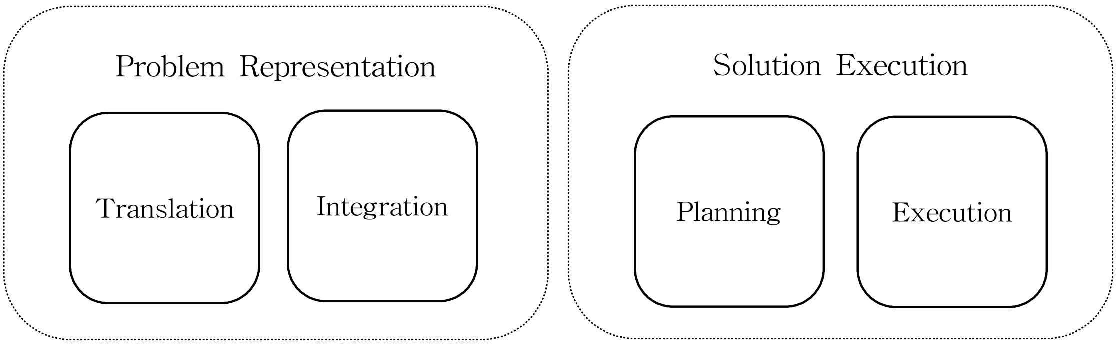
Particularly, ADHD children showed low achievement in word problems (Re, Lovero, Cornoldi, & Passolunghi, 2016). Whereas calculation problems are already set for computation, word problems require students to determine important information among words and numbers and plan out solutions. ADHD children appeared to remember literal and numerical information from the word problems but confused by them in the problem-solving process. This implies that attention plays a role in the success of word problem-solving.

Mayer (1980) has proposed the model of mathematical problem solving (Figure 1) consisted of two major components: problem representation and solution execution. There are two phases in the problem representation component. The first is the translation phase. In this phase, a problem solver reads a word problem and constructs an internal representation of the problem. After this literal phase, the problem solver moves to the integration phase. In order to understand the meaning of the problem, the solver uses the schema to judge which information is important and which is not. Thus, the problem solver integrates the problem and schema such as cause/change, combine, and compare. The third phase, which is a part of the solution execution component, is where the problem solver plans out for the types of operation and sequence of using them to solve the problem. Finally, in the fourth phase, also a part of the solution execution component, the problem solver completes the computations using algorithmic knowledge. These four phases generally take place in order. However, the problem solver always goes back and forth self-monitoring when needed (Mayer, 1980).

In particular, the problem representation component requires attention. In this phase, a student reads a word problem and construct an internal representation using linguistic knowledge and factual knowledge. This process involves copying the problem word for word, substituting terms like “two times” for twice, and locating words or phrases describing variables (Mayer, 1980). The relevant variables and their relations are picked up and integrated with problem solvers’ schematic knowledge to form a problem structure for later mathematical operation (Kingsdorf & Krawec, 2014). Even though strategic knowledge is well-established and computation skills are automatized, successful performance is not guaranteed without accurate problem representation. Therefore, the problem representation component demands attention in selecting only relevant variables and understanding their relations so that problem solver can move on to the next phase with correct representation.

Figure 1

Mayer's Model of the Mathematical Problem-Solving



Linguistic structure of word problems contributes to problem representation (Song, 1992). Especially, relational terms in word problems that are used to define a variable in terms of another bring difficulty in representing the problem correctly (Lewis & Mayer, 1987). For example, in Question 1 shown in Table 2, the relational term "more than" is used to define the variable of the unknown variable-the number of Tom's marbles in relation to the other variable, i.e., the number of Joe's marbles. In Question 4, the relational term "less than" is used to determine the unknown variable-the number of Tom's marbles. During the problem representation stage, the problem solver reads the relational terms in the word problems, integrate the information to a schema related to the terms, and form an internal representation of the problem.

However, if the relational terms prime an incorrect schema for the problem solution, a correct answer will not be yielded regardless of precise computational skills. In Question 5, for example, the relational term “less than” can be considered as a cue activating the schema related to subtraction, which leads to an incorrect outcome. Word problems with relational terms that are stated inconsistently with the required arithmetic operation are called inconsistent language problems (Lewis & Mayer, 1987). Question 4 in Table 4 is an example. The question includes the relational term “less than” which is inconsistent with the required operation, addition. On the other hand, the relational term “more” in Question 1 is consistent with the required arithmetic operation, addition. This form of problem is referred to as consistent language problems (Lewis & Mayer, 1987). These two forms of problems bring out different levels of difficulty in problem-solving. The difficulty in processing inconsistent language while solving word problems has been reported. Not only children but even adults showed a higher error rate on inconsistent language problems than on consistent language problems (Lewis & Mayer, 1987). In addition, students with high accuracy on inconsistent problems appeared to spend a longer time re-reading the problems (Hegarty, Mayer, & Green, 1992).

The quality of the information provided in word problems contributes to math word problem-solving performances (Daroczy, Wolska, Meurers, & Nuerk, 2015). Containing irrelevant linguistic and numerical distracters increase the complexity of problems, thus demanding more cognitive resources. It was found that word problems with irrelevant information resulted in a higher error rate (Kim & Kim, 2006; Muth, 1992). Thus, the existence of extra information requires problem solvers to process more information and to employ additional cognitive steps.

Several individual cognitive factors that influence the translation and integration phases have been identified. For example, learning styles affect the problem representation. Song (1992) studied the role of learning style in math word problem recalling among college students. He found out that students with surface and reiterative learning style made more errors than those with deep-elaborative processing style. Additionally, he found out that the error types differed among students. The deep and elaborative processing learners made simple recalling errors, while the surface and reiterative processing learners appeared to change the problem structure and recall the problems in a contradicting and incomplete manner.

The role of attention in the understanding phase of word problems has been rarely investigated. Students who pay attention to only keywords for planning computational operations and do not deeply process all the contextual information to understand the word problems could not represent the problem accurately, thus resulting in errors. The literature has found out that distinctive font styles could enhance students′ alerted state (Kercood, Zentall, Vinh, & Tom-Wright, 2012). Even a “+” sign evokes brain regions associated with spatial attention orienting (Mathieu et al., 2018). Children with attention deficits prefer salient features of stimuli. Therefore they are expected to have great difficulty focusing on plain and subtle but relevant stimuli (Kercood, Zentall, & Lee., 2004). However, it is unclear how attention plays a role in the word problem-solving process.

The attention-related scores yielded from the CANT might explain the individual differences in the performance of mathematical word problem-solving. If the CANT wants to be considered ecologically valid, the experimental situation, sampled responses, and stimuli used in the test should represent those in the mathematical word problem-solving. The accuracy performance in the CANT should be generalized to that in the mathematical word problem-solving. The response time extracted from the CANT should be able to represent that in the mathematical problem-solving. Besides, the performance pattern of speed and accuracy shown in the CANT should be replicated in the mathematical problem-solving.

Furthermore, attention networks should reveal their distinctive role in different types of word problems. Alerting means to generate and maintain an alerted state so that one can locate and respond to stimuli. If a student

Table 2

Examples of Comparing Word Problems

Language Types	Example of Word Problems	Required Computational Operation	Unnecessary Information
Consistent	Question 1. Joe has 3 marbles. Tom has 5 *more* marbles *than* Joe. How many marbles does Tom have?	Addition	Not Included
Consistent	Question 2. Joe has 3 marbles. Jane has 5 marbles. Tom has 5 *more* marbles *than* Joe. How many marbles does Tom have?	Addition	Included
Consistent	Question 3. Joe has 8 marbles. Tom has 5 marbles *less than* Joe. How many marble does Tom have?	Subtraction	Not Included
Inconsistent	Question 4. Joe has 3 marbles. He has 5 marbles *less than* Tom. How many marbles does Tom have?	Addition	Not Included
Inconsistent	Question 5. Joe has 3 marbles. He has 5 marbles *less than* Tom. He has 4 *more* marbles *than* Jane. How many marbles does Tom have?	Addition	Included

Note. The examples of word problems were extracted from Lewis & Mayer (1897). The relational terms in the word problems are italicized. The unnecessary information for correct problem representation is underlined.

stays less alerted in a regular situation and can not efficiently change to an alerted state unless a warning signal is preceded, this student would yield a high alerting score in the CANT. Therefore, students with a higher alerting score would show lower performance in a set of word problems written in plain fonts, that is, not highlighted nor underlined. Orienting refers to locate and shift attention to the important stimuli. Students with a higher orienting score can not orient stimuli in space efficiently unless a cue that indicates the target stimuli' location is presented. Therefore, it could be assumed that the student with a higher orienting score would show lower performance in word problems in which the relevant variables are not clearly presented. Executive control refers to monitor and dissolve conflicts that were generated by surrounding distracters. Students with a higher executive control score can be more distracted by inconsistent flankers around the target. Therefore, it could be assumed that the student with a higher executive control score would show lower performance in word problems in which inconsistent relational terms are used to describe the unknown variables. If these CANT scores can be generalized to mathematical word problem-solving performances, the CANT could be determined as an ecologically sound attention measure.

Ⅲ. Study 1

1. Research Question

Study 1 aims to test the psychometric properties of the CANT with Korean samples, which have a different cultural background from those the test originally targeted. To achieve this goal, the item quality, reliability, and validity of the CANT in Korean samples were examined. The item quality was tested in terms of its discrimination and the validity of cue and flanker manipulations at the item level. The reliability was investigated in terms of its internal consistency and stability over time. The validity was investigated in terms of its concurrent validity with other attention measures.

Research Question:
How valid is the CANT psychometrically in the setting of Korean elementary school?

Sub-questions:
1. How relevant are the items consisting of the CANT?
2. How reliable is the CANT?
3. How valid is the CANT?

2. Methodology

1) Participant

The target population for this study included elementary school students in a province, Korea. Student participants were recruited in an elementary school located in the metropolitan area. Since there were multiple classes per grade, one class from each grade was randomly selected, thus six classes in total. The researcher visited the homerooms of the selected classes and explained the study's purpose, the tests to take, and rewards. Then consent letters were distributed to those who were willing to participate in the study. They were asked to return the letter with their parents or legal guardians signatures within a week. One hundred and fifty three consent letters were distributed, and 71 of them returned.

Other than student participants, parents and homeroom teachers also participated in this study. The parents of the fourth to sixth grades students who wished to participate were asked to participate in the study. Thirty-one parents agreed to participate and returned their consent letters via their children. The homeroom teachers of the selected classes in grades four to six were asked to participate. Since one class per each grade was randomly selected to recruit student participants in this study, three homeroom teachers of the selected classes were asked to participate. All of them agreed to be part of the study.

However, different numbers of student participants were employed in answering the research questions. Seventy-one participants' responses were used for item analysis, the split-half reliability test and calculating Cronbach's alpha. Thirty-five students' data from grades four to six were analyzed in the test-retest reliability test and concurrent validity test. Furthermore, the data cleaning process revealed some missing data and outliers, which were excluded in the following statistical analysis. Hence the actual sample sizes for each analysis could be different from the numbers stated above. Table 3 shows the demographic characteristics of participants.

Table 3

Demographic Characteristics of the Participants

Grade	Student		Parent		Teacher	
	n	%	*n*	%	*n*	%
1	10[a]	14.0	-	-	-	-
2	10[a]	14.0	-	-	-	-
3	16[a]	23.0	-	-	-	-
4	13[ab]	18.0	12	38.7	1	33.3
5	11[ab]	15.5	9	29.0	1	33.3
6	11[ab]	15.5	10	32.3	1	33.3
Total	71[ab]	100	31	100	3	100

Note. The parents of the fourth to sixth-grade students who wished to participated were asked to participate in this study. The homeroom teachers of the selected class in grades four to six were asked to participate in this study. N^a = the student participants for item analysis, split-half reliability test and Cronbach's alpha estimation; N^{ab} = the student participants for test-retest reliability test and concurrent validity test as well as item analysis and split-half reliability test.

2) Measurement

(1) The Child Attention Network Test (CANT)

Rueda et al. (2004) revised the ANT to develop a more child-friendly version, the CANT. They changed the stimuli from a set of arrows to an array of yellow fish. They created a story - feeding a right fish, and added a clear feedback sound in the practice block. These modifications were to motivate young subjects to respond to the target. It consisted of an introduction part, a practice block, and three test blocks. Subjects were asked to read the introduction to figure out how to take the test and participate in the practice block and later test blocks. Subjects were asked to respond to the stimuli described in detail below, as quickly and accurately as possible, by judging the target stimulus's direction and pressing the designated buttons on the computer keyboard. There were 48 trials in a block and 144 trials in total in the test. The practice block consisted of 24 trials. As Rueda and her colleagues (2004) suggested, three networks' efficiency is calculated individually using the median respond time (RT_{mdn}) for correct responses.

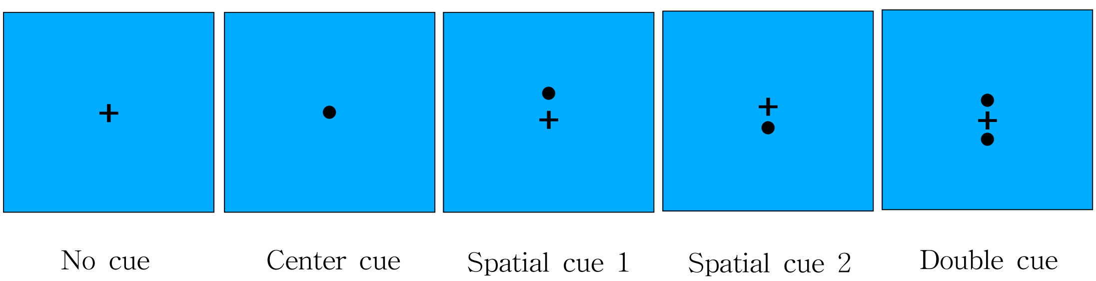
In this study, the format of the CANT was replicated with some modifications in the practice block. The CANT was developed for English speaking users. Therefore, language translation was needed in the introductory part of the test. In addition, the feedback sound was replaced since the original sound in the CANT, "Woohoo!" for correct response and a low beep sound for incorrect response, could be culturally unfamiliar for Korean children. A new bell sound typically used in Korea for correct answers and a beep sound for incorrect answers replaced the original feedback sounds. Other than these modifications, the original form of the CANT was maintained in this study.

The stimuli used in the CANT were yellow fish with an arrow in the body, pointing in the same direction of where the fish was heading. Sometimes the fish appeared alone, and other times with two more same-looking fish on its right and left sides. The stimuli were presented against the light blue background. When the test began, a fixation mark (+) appeared in the center of the background, and later the target was presented above or below the fixation according to cue and flanker condition. The stimuli are shown in Figure 2.

Stimuli were preceded by a cue shaped like a black dot (●). The cue was to signal subjects that an array of fish was about to show up and provide information on where the stimuli would appear. There were four cue types: no cue, center cue, double cue, and spatial cue. On the no cue condition, no cue appeared before stimuli giving no information of when and where the stimuli would appear. On the center cue condition, a black dot appeared at the center of the screen replacing the fixation mark and signaling appearance of the stimuli. On the double cue condition, a black dot appeared above and below the fixation mark at the same time signaling appearance of the stimuli, but without information of their location. On the spatial cue condition, a black dot was presented either above or below the fixation mark providing information both of when and where the stimuli would appear. The cue conditions are described in Figure 2.

Figure 2

Cue Conditions in the CANT



As Figure 3 shows, there were three types of flankers: congruent, incongruent, and neutral types. Two flankers flanked the target stimulus on the left and right side of it in the array. Congruent flankers were the same yellow fish pointing in the same direction as the target fish. Incongruent flankers pointed in the opposite direction from the target fish. In trials of neutral flanker, no fish appeared except the target fish. In total, there were four cue types and three flanker types in the CANT. The target stimulus was presented based on the combination of the cue and flanker types, which yielded twelve trials comprising a cue-flanker combination set. Forty-eight trials were offered in a block randomly.

As Figure 4 illustrates, each trial consisted of a fixation period, a cue presentation when applicable, another fixation period, a target and flanker period, and a post-target fixation period. During the fixation period, the fixation mark (+) was presented for a random variable duration (400-1600 msec). Then a cue (●) appeared according to the cue type of the trial for 150 msec, followed by another fixation period for 450 msec. After the second fixation period, the target and flankers were presented until the subject made a response, but no longer than 1700 msec. Finally, the post-target fixation mark appeared for 1000 msec. After the last fixation period, the next trial began.

The estimates of the CANT were based on the response time and accuracy data. The three attentional networks' efficiency was calculated using the

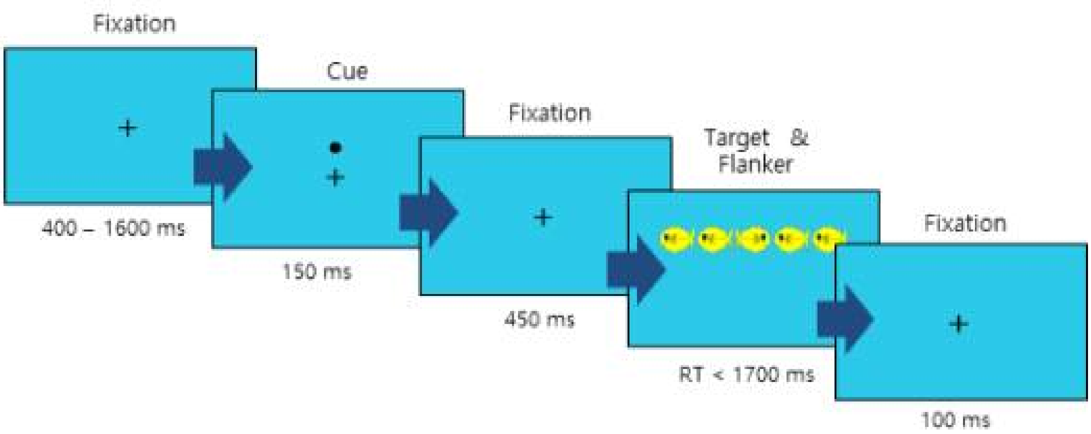
Figure 3

Flanker Conditions in the CANT

Congruent flanker condition | Incongruent flanker condition | Neutral flanker condition median response time of accurate trials of each condition. The CANT alerting score (CANT_A) was calculated by subtracting the median RT of double cue condition from the median RT of no-cue condition across all the flanker conditions. The larger CANT_A is, the more difficult it is to create and maintain an alerted state without a warning cue. On the other hand, it could reflect more efficient use of warning cues. The CANT orienting score (CANT_O) was calculated by subtracting the median RT of the spatial cue condition from the median RT of the central cue condition. The larger CANT_O is, the more difficult it is to disengage attention from the location where initial attention was oriented. However, this also can be interpreted as an efficient use of spatial cues in the orienting network. The CANT executive control score (CANT_EC) was calculated by subtracting the median RT of congruent flanker conditions from incongruent flanker conditions across all the cue conditions. The larger CANT_EC is, the more difficult it is to resolve conflict (Fan et al., 2002; MacLeod et al., 2010; Rueda, et al., 2004). The median RT of correct responses in the entire CANT ($CANT\_CRT_{mdn}$) and the percentage of correct responses of the entire CANT (CANT_PCR) were also calculated.

Figure 4

The Trial Phase of the CANT



(2) The Frankfurt Adaptive Concentration Test-Ⅱ (FACT-Ⅱ)

The FACT-Ⅱ was a performance-based measurement assessing attention, originally developed by Moosbrugger and Goldhammer (2014) and normalized in Korea by Oh (2018). Subjects was required to discriminate between the target and the non-targets presented on the computer screen. The stimuli are geometrical figures of shapes and dots within them. The target stimuli were a circle with three dots in it and a square with two dots in it, whereas a circle with two dots and a square with three dots were non-targets. FACT-Ⅱ had three subtests: a test with each item provided individually (FACT-Ⅱ-E), a test with item series provided (FACT-Ⅱ-S), and a test with item series provided with record of reaction time (FACT-Ⅱ-SR). Among these subtests, the FACT-Ⅱ-SR was used in this study since it has been standardized among Korean elementary students. In the FACT-Ⅱ-SR, a row of 10 items, as shown in Figure 4, appeared on the screen, and subjects responded to the items which were marked by an arrow by pressing the "1" key on the keyboard for the targets and the "0" key for non-targets as fast as possible. When responding to a row of items was completed, another row of 10 items appeared. The test finished automatically after six minutes has passed. After a one-minute long practice block, the actual FACT-Ⅱ-SR began.

Based on the response time and the number of correct response, three estimates were calculated. The power of concentration performance (FACT-Ⅱ

Figure 4

The Stimuli for FACT-Ⅱ-SR

-SR_CP) was calculated using RT for accurate responses, as appears in Equation 1. The higher FACT-Ⅱ-SR_CP index indicates that the participant concentrates at a faster speed. As Equation 2 shows, concentration's correctness (FACT-Ⅱ-SR_CC) was calculated as a ratio of the number of accurate responses and the total number of responses. A higher FACT-Ⅱ-SR_CC score indicates higher correctness in the task. Concentration balance (FACT-Ⅱ-SR_CB) was calculated using the standard deviation of item serve time, as shown in Equation 3. The higher FACT-Ⅱ-SR_CB score indicates the consistency of attention during the test. The split-half reliability of the FACT-Ⅱ-SR test was reported as .96 for FACT-Ⅱ-SR_CP, .88 for FACT-Ⅱ-SR_CC, and .88 for FACT-Ⅱ-SR_CB. The test-retest reliability was .75 for FACT-Ⅱ-SR_CP, .65 for FACT-Ⅱ-SR_CC, and .72 for FACT-Ⅱ-SR_CB.

$$\text{FACT-II-SR\_CP} = \frac{W - 2 \times F}{\sum_{i=m+1}^{n} t_1} \times 100{,}000. \qquad (1)$$

W= the number of responded items (n-m)

n= the last item number responded

m= the last item number responded during the practice block

t_1= the total response time until the i^{th} item (msec)

F= the number of incorrect response from (m+1) item to n

$$\text{FACT-II-SR\_CC} = \frac{W - F}{W} \times 100. \qquad (2)$$

W = the number of responded items

n = the last item number responded

m = the last item number responded during the practice block

F = the number of incorrect response from (m+1) item to n

$$\text{FACT-II-SR\_CB} = \log(1 + \frac{1}{S_b}) \times 10{,}000. \quad (3)$$

$$S_b^2 = S^2 - S_t^2$$

$$S^2 = \frac{\sum_{i=m+1}^{n} t_1 (t_1 - AM)^2}{W}$$

$$AM = \frac{\sum_{i=m+1}^{n} t_1}{W}$$

S^2 = variation of item serve time

S_b = deviation of response time

S_t^2 =technical variation conditioned by computer algorithm

(3) Korean Attention Deficit Hyperactivity Disorder Diagnostic Scale (KADHDDS)

KADHDDS was developed based on the ADHD diagnostic criteria in DSM-Ⅳ-TR and Gilliam Attention Deficit Hyperactivity Disorder Test, and normalized among Korean children and young adults (Yoon & Lee, 2004). It was used for the parents of the student participants to evaluate their childrens' attention-related behaviors. The scale included three subscales (Inattention, Hyperactivity, and Impulsivity), consisting of 13 items, 10 items, and 10 items for each subscale. The items for Inattention was used for this study. Based on the severity of the symptoms, parents rated their children's behaviors on a scale of 0 to 2, where 0 indicated the absence of symptoms, 1 indicated mild symptoms, and 2 indicated severe symptoms: the higher score, the more severe symptoms of inattentiveness. Cronbach's α of the Inattention subscale was .90, and test-retest reliability was .90 (Ji, 2019).

(4) Korean Teacher ADHD Rating Scale (K-ADHDRS-T)

The K-ADHDRS-T was the Korean version of the ADHD Rating Scale-Ⅳ

(School Version), which was originally developed by DuPaul et al. (1997). It was first validated in Korea by So et al. (2002) and normalized by Kim et al. (2003). The K-ADHDRS-T was used for the teacher participants to evaluate their students' attention-related behaviors. The K-ADHDRS-T was an 18-item questionnaire composed of two subscales, which were Inattention and Hyperactivity-Impulsivity with nine items, respectively. Only items from the Inattention subscale were used for this study. The items asked the frequency of problematic behaviors of children related to inattentiveness. Teachers were asked to rate the frequency of behaviors related to inattentiveness on a 0 to 3 scale: 0 = never or rarely; 1 = sometimes; 2 = often; 3 = very often. The Cronbach's α of the scale ranged from .77 to .89 (So et al., 2002).

3) Data Collection

The data were collected for three months, starting from June to August of 2020. Subjects took the CANT test in a classroom of the school they were attending. The subjects sat at a desk on which a laptop computer was located. During the practice block, the researcher monitored their performance to see whether they understood the test. Feedback on the accuracy of their response was given during the practice block; a bell sound for accurate response and a beep sound for incorrect response. After 24 trials of practice, the subjects were asked to rest and start the real test when ready. While they were taking the experimental blocks, no feedback was given. The researcher stayed away from them not to disturb, but still was in the same room. The subjects could rest as long as they wanted between the experimental blocks, and they proceeded to the next block by pressing the spacebar. It took approximately 20 minutes to complete the test, including the practice block. Subjects from grades one to three received a pack of snack as a reward.

Thirty-five students in grades four to six took the FACT-Ⅱ-SR test after taking the CANT. The researcher explained how to respond to the stimulus, and then the subjects took a practice test approximately for three minutes. When their response was incorrect, feedback was given from the computer correcting their response. During the real test, no feedback was provided. It took approximately nine minutes to complete the FACT-Ⅱ-SR, including the practice trials. These students revisited the classroom after 10-14 days from the first CANT test. They repeated the procedure except taking the FACT-Ⅱ-SR. They received a 5,000-won gift certificate as a reward.

The parents of thirty-one students in grades four to six evaluated their children′s attention-related behaviors on the KADHDDS questionnaire, which was sent via their children. Their teachers were asked to rate the students′ attention-related behaviors on the K-ADHDRS-T questionnaire, as well. The questionnaires took approximately three minutes to complete. Their children returned the parents′ questionnaires to the researcher. The K-ADHDRS-T questionnaires were delivered to the teachers by the researcher in the last week of the semester, and the researcher visited the subjects' classrooms to collect the teachers' questionnaires.

The CANT was presented via Inquisit 5, which is a commercial software package for generating and running psychological experiments on a computer. The stimuli were presented on the screen of the laptop, and subjects registered their responses by pressing two designated keys on the keyboard, 'E' for leftward direction and 'I' for the rightward direction of the target stimuli. In order to minimize incorrect response caused by confusion in direction, a sticker on which "왼" meaning leftward in Korean and "←" were marked was attached on the "E" key, and "오" meaning rightward in Korean and "→" for the "I" key. The FACT-Ⅱ-SR was presented on the website of INPSYT, a commercial website for purchasing and running psychological tests. On the same laptop used for the CANT, subjects viewed the screen where the FACT-Ⅱ-SR stimuli were presented. They used the keys of "1" for correct answers and "0" for incorrect answers to register their responses.

4) Data Analysis

The quality of items was tested by applying item discrimination and item validity. Item analysis is a statistical analysis to evaluate tests based on the quality of individual items, items sets, and entire sets of item (McCowan & McCowan, 1999: 3). The CANT comprised three blocks containing 48 trials in each. Each trial was constructed by combining one of the four cue types and one of the three flanker types, yielding 12 different cue × flanker combinations. Since these twelve combinations were presented four times repetitively and randomly in a block, each cue and flanker condition was considered as a unit for item analysis.

Item discrimination was tested by examining the correlation coefficient of the mean of the median RT with the median RT of the overall test and the correlation coefficient of the percentage of correct responses of each item with the percentage of correct responses of the overall test. The significant coefficient indicates that the item contributes to the test discrimination (Vorstenbosch et al., 2013). Item validity was assessed by comparing the median response time and the percentage of correct responses depending on the cue and flanker condition of each item. ANOVA was employed for the test.

The reliability of the CANT was assessed in terms of its internal consistency and stability. Internal consistency was evaluated by using Cronbach's alpha and split-half reliability. Since the CANT items were 12 combinations of cue and flanker conditions, the percentage of correct responses and median response time of each item were correlated with those of the test when calculating Cronbach's alpha. For the split-half reliability test, Pearson correlation coefficients between the scores in three blocks were calculated. In order to see the test-retest reliability of the CANT, the intraclass correlation coefficient (ICC) between the results at two different times was estimated. ICC is a widely used reliability index in test-retest reliability analysis (Koo & Li, 2016). Since test-retest reliability demonstrates the variation in estimates taken by the same instrument on the same subjects, the two-way mixed-effects model and absolute agreement definition were selected during the analysis. SPSS calculates ICC based on the formulation suggested by McGraw and Wong (1996), as shown in Equation 4 (for terminology in the equation, refer to McGraw & Wong's study). SPSS was used for the analysis.

$$ICC = \frac{MS_R - MS_E}{MS_R + (k-1)MS_E + \frac{k}{n}(MS_C - MS_E).} \quad (4)$$

MSR = mean square for rows

MSE = mean square for error

MSC = mean square for columns

n = number of subjects

k = number of measurements

The validity of the CANT was assessed by calculating Pearson correlation coefficients between the scores obtained in the CANT and the criterion measures, i.e., FACT-Ⅱ-SR, KADHDDS, and K-ADHDRS-T. SPSS was used for the analysis.

3. Results

Study 1 examined psychometric properties of the CANT in Korean samples. The results of the tests are as follows.

1) Item Quality

Item analysis was conducted in order to see the contribution of each item to the discrimination and validity of the CANT, As Table 4 indicates, there were variations in CANT_PCR and $CANT\_RT_{mdn}$ among the cue × flanker conditions. The mean of CANT_PCR ranged from 93.3 (*SD* = 10.1) to 95.7 (*SD* = 7.8) in the no cue condition, and from 94.7 (*SD* = 8.1) to 96.7 (*SD* = 7.6) in the spatial cue condition. Meanwhile, the mean of CANT_PCR ranged from 95.7 (*SD* = 7.8) to 97.4 (*SD* = 7.0) in the congruent flanker condition, and from 93.3 (*SD* = 10.1) to 94.7 (*SD* = 8.1) in the incongruent flanker condition.

The mean of $CANT\_RT_{mdn}$ also varied among conditions. The mean of $CANT\_RT_{mdn}$ in the no cue conditions ranged 642.44 milliseconds (*SD* = 140.40) to 726.44 milliseconds (*SD* = 159.30), whereas that in the spatial cue conditions ranged 560.90 milliseconds (*SD* = 121.55) to 655.64 milliseconds (*SD* = 152.83). The mean of $CANT\_RT_{mdn}$ ranged from 570.56 milliseconds (*SD* = 125.33) to 665.38 milliseconds (*SD* = 154.06) in the congruent flanker conditions and ranged from 655.64 milliseconds (*SD* = 152.83) to 726.44 milliseconds (*SD* = 159.30) in the congruent flanker conditions.

Table 4

Descriptive Statistics for CANT_PCR and CANT_RT$_{mdn}$ by Condition (N = 71)

Condition		CANT_PCR (%)		CANT_RT$_{mdn}$ (msec)	
Cue	Flanker	*M*	*SD*	*M*	*SD*
No	Congruent	95.7	7.8	665.36	154.06
	Incongruent	93.3	10.1	726.44	159.30
	Neutral	94.8	7.2	642.44	140.40
Center	Congruent	95.8	10.0	610.90	129.17
	Incongruent	93.7	9.2	698.85	138.19
	Neutral	94.1	9.2	594.75	129.15
Double	Congruent	97.4	7.00	596.75	120.99
	Incongruent	94.5	8.3	684.73	148.76
	Neutral	95.5	6.4	598.89	134.99
Spatial	Congruent	96.7	7.7	570.56	125.33
	Incongruent	94.7	8.1	655.64	152.83
	Neutral	96.5	6.7	560.90	121.55

Note. CANT_PCR = the percentage of correct responses in the CANT; CANT_RT$_{mdn}$ = the median response time in the CANT.

IItem-total correlation analysis was conducted to see how well each item contributes to the test result. As Table 5 demonstrates, the item-total correlation for items in the no cue condition ranged from .34 to .80, .41 to .80 in the center cue condition, .45 to .76 in the double cue condition, and .60 to .77 in the spatial cue condition. The item-total correlation for items in the congruent flanker conditions ranged between .65 and .80, those in incongruent flanker conditions ranged between .41 to .77. and those in neutral flanker conditions ranged between .34 to .74. None of the conditions appeared to contribute to improving Cronbach's α when deleted. The item-total correlation values between .20 and .39 are considered good discrimination and the value of .40 and above to be very good (McCowan & McCowan, 1999). Therefore, the accuracy of each item could be determined to show good and very good discrimination.

As Table 5 demonstrates, all the means of the median RT of each item showed a high correlation with the total median response time. The item-total correlation ranged from .91 to .92 in the no cue conditions, .92 in the center cue condition, .89 to .93 in the double cue condition, and .90 to .94 in the spatial cue condition. The item-total correlation ranged from .92 to .94 in the congruent flanker condition, .91 to .94 in the incongruent flanker condition, and .89 to .92 in the neutral flanker condition. None of the conditions appeared to contribute to improving α when deleted, considerably. Therefore, the items of each condition could be determined to show very good discrimination in terms of RT_{mdn}.

Table 5

Result of Item-total Correlation Analysis of PCR and RT_{mdn} by Condition (N = 71)

Condition		CANT_PCR_item		CANT_RT_{mdn}_item	
Cue	Flanker	Item-Total Correlation	Cronbach's α If Item Deleted	Item-Total Correlation	Cronbach's α If Item Deleted
No	Congruent	.65	.89	.92	.98
	Incongruent	.52	.89	.91	.98
	Neutral	.34	.90	.91	.98
Center	Congruent	.80	.88	.92	.98
	Incongruent	.41	.90	.92	.98
	Neutral	.74	.88	.92	.98
Double	Congruent	.72	.88	.93	.98
	Incongruent	.63	.89	.91	.98
	Neutral	.45	.90	.89	.98
Spatial	Congruent	.76	.88	.94	.98
	Incongruent	.77	.88	.94	.98
	Neutral	.60	.89	.90	.98

Note. CANT_PCR_item = the percentage of correct responses for each item in the CANT; CANT_RT_{mdn}_item = the median response time for each item in the CANT.

The effects of cue and flanker manipulation on the accuracy and speed of each item were examined. Table 6 shows the result of ANOVA with the percentage of correct responses of each item (CANT_PCR_item). There was a significant main effect of flanker condition on the accuracy, $F(2, 213) = 5.78$, $p = .00$. No significant main effect was found for the cue condition, $F(3, 284) = 1.87$, $p = .13$. As Table 7 demonstrates, the post-hoc test revealed a higher percentage of correct responses in the congruent flanker condition than in the incongruent condition. No significant difference was found in the accuracy between the congruent and neutral conditions.

The effects of cue and flanker condition on the RT_{mdn} of each item ($CANT\_RT_{mdn}\_item$) was also examined. Table 6 shows the result of ANOVA of four cue conditions × three flanker types using the RT_{mdn} data from 71 subjects. The result demonstrated a significant main effects of cue condition, $F(3, 284) = 12.82$, $p = .001$, and flanker condition, $F(2, 213) = 37.31$, $p = .001$. However, there was no significant interaction between cue condition and flanker condition. As Table 8 demonstrates, the RT_{mdn} of no cue condition was significantly longer than those of center cue, double cue, and spatial cue conditions. The RT_{mdn} of the spatial cue condition was significantly shorter than those of the center cue and no cue condition. However, the RT_{mdn} of the spatial cue condition and center cue condition did not show a significant difference. The RT_{mdn} with incongruent flanker type was significantly longer than those with congruent and neutral flanker types. The RT_{mdn} with the congruent flanker type was not significantly different from that with the neutral flanker.

The results demonstrated that the flanker condition affected not only the accuracy but also the response time of each item. Cue manipulation appeared to have an impact on the response time of each item. To sum up, the flanker manipulation could be determined to be valid in terms of its effects on the accuracy and response time of each item. The cue manipulation could be determined to be valid in terms of its effect on the response time of each item.

Table 6

Result of 4 × 3 (Cue × Flanker) ANOVA on CANT_PCR_item and CANT_RT_{mdn}_item (N = 71)

	Source	Type Ⅲ Sum of Squares	*df*	Mean Square	*F*	*p*	Partial eta Squared
CANT_PCR_item	Corrected model	1238.92	11	112.63	1.66	.08	.02
	Intercept	7726077.14	1	7726077.14	114157.29	.00	.99
	Cue	380.15	3	126.72	1.87	.13	.01
	Flanker	782.64	2	391.32	5.78	.00	.01
	Cue×Flanker	76.13	6	12.69	.19	.98	.00
	Error	56850.55	840	67.68			
	Total	7784166.60	852				
	Corrected total	58089.46	851				
CANT_RT_{mdn}_item	Corrected model	2194559.20	11	199505.38	10.40	.00	.12
	Intercept	342306762.63	1	342306762.63	17850.87	.00	.96
	Cue	737556.04	3	245852.01	12.82	.00	.04
	Flanker	1430823.17	2	715411.59	37.31	.00	.08
	Cue×Flanker	26179.99	6	4363.33	.23	.97	.00
	Error	16107767.17	840	19175.91			
	Total	360609089.00	852				
	Corrected total	18302326.37	851				

Note. CANT_PCR_item = the percentage of correct responses for each item in the CANT; CANT_RT_{mdn}_item = the median response time of each item in the CANT.

Table 7

Result of Scheffé Test on CANT_PCR_item (N = 71)

Condition			Mean difference (I-J)	Std. error	*p*	Sheffé
	(I) Cue	(J) Cue				
Cue	Center	Double	−1.29	.80	.45	−
		No	−.08	.80	1.00	
		Spatial	−1.45	.80	.35	
	Double	Center	1.29	.80	.45	
		No	1.21	.80	.51	
		Spatial	−.16	.80	1.00	
	No	Center	.08	.80	1.00	
		Double	−1.21	.80	.51	
		Spatial	−1.37	.80	.40	
	Spatial	Center	1.45	.80	.35	
		Double	.16	.80	1.00	
		No	1.37	.80	.40	
	(I) Flanker	(J) Flanker				
Flanker	Congruent	Incongruent	2.35	.69	.00	Congruent > Incongruent
		Neutral	1.14	.69	.25	
	Incongruent	Congruent	−2.35	.69	.00	
		Neutral	−1.20	.69	.22	
	Neutral	Congruent	−1.14	.69	.25	
		Incongruent	1.20	.69	.22	

Note. CANT_PCR_item = the percentage of correct responses for each item in the CANT.

Table 8

Result of Scheffé Test after ANOVA on CANT_RT_{mdn}_item (N = 71)

Condition			Mean difference (I-J)	Std. error	*p*	Sheffé
Cue	(I) Cue	(J) Cue				No > Center, No > Double, No > Spatial, Center > Spatial
	Center	Double	8.04	13.42	.95	
		No	-43.25	13.42	.02	
		Spatial	39.13	13.42	.04	
	Double	Center	-8.04	13.42	.95	
		No	-51.29	13.42	.00	
		Spatial	31.09	13.42	.15	
	No	Center	43.25	13.42	.02	
		Double	51.29	13.42	.00	
		Spatial	82.38	13.42	.00	
	Spatial	Center	-39.13	13.42	.04	
		Double	-31.09	13.42	.15	
		No	-82.38	13.42	.00	
Flanker	(I) Flanker	(J) Flanker				Incongruent > Congruent, Incongruent > Neutral
	Congruent	Incongruent	-80.52	11.62	.00	
		Neutral	11.64	11.62	.61	
	Incongruent	Congruent	80.52	11.62	.00	
		Neutral	92.17	11.62	.00	
	Neutral	Congruent	-11.64	11.62	.61	
		Incongruent	-92.17	11.62	.00	

Note. CANT_RT_{mdn}_item = the median response time for each item in the CANT.

2) Reliability of the CANT

Reliability of the CANT was assessed in terms of its internal consistency and stability of the test results. Cronbach's alpha for CANT_PCR appeared to be .90 and .99 for CANT_RT_{mdn}. Table 9 demonstrates the split-half reliability of the CANT. Because the CANT consists of three blocks, the scores in different blocks were compared to test the consistency of the results across blocks. CANT_A showed coefficients of .16 to .26, CANT_O appeared to show coefficients of -.17 to .16, and CANT_EC showed coefficients of .08 to .39 between three blocks. CANT_PCR showed coefficients of .52 to .65, and CANT_RT_{mdn} showed coefficients of .93 to .94 between blocks. CANT_RT_{mdn} appeared consistent across three blocks and CANT_PCR appeared moderately consistent. However, the other CANT scores demonstrated a low consistency across the blocks.

Table 9

Result of Pearson Corelation Analysis between the Scores in the Three Blocks of the CANT (N = 71)

	Block 1 & 2	Block 1 & 3	Block 2 & 3
CANT_A	.26	.16	.21
CANT_O	.16	-.17	.05
CANT_EC	.39	.37	.08
CANT-PCR	.65	.53	.52
CANT_RT_{mdn}	.93	.93	.94

Note. CANT_A = the CANT alerting score; CANT_O = the CANT orienting score; CANT_EC = the CANT executive control score; CANT_PCR = the percentage of correct responses in the CANT; CANT_RT_{mdn} = the median response time in the CANT.

In order to see the test results' stability over time, test-retest analysis was conducted. Thirty-five students' data in grades four to six were collected for this analysis. However, due to missing data and outliers, 30 students' data were analyzed. Table 10 shows the result of intraclass correlation (ICC) analysis. The ICC was .46 for CANT_A, .16 for CANT_O, .40 for CANT_EC, .17 for CANT_PCR, and .90 for $CANT\_RT_{mdn}$. The ICC coefficients of CANT_A, and $CANT\_RT_{mdn}$ were statistically significant, and CANT_EC tended to have a weak correlation between the scores measured at two different times. The coefficients for CANT_O and CANT_PCR were not statistically significant.

Table 10

Result of Intraclass Correlation Analysis of the CANT Scores (n = 30)

Variable	ICC	95% Confidence Interval		p
		Lower Bound	Upper Bound	
CANT_A	$.46^{*}$	-.09	.74	.04
CANT_O	.16	-.77	.60	.32
CANT_EC	$.40^{\dagger}$	-.18	.70	.07
CANT_PCR	.17	-.59	.58	.29
$CANT\_RT_{mdn}$	$.90^{***}$	.40	.97	.00

Note. CANT_A = the CANT alerting score; CANT_O = the CANT orienting score; CANT_EC = the CANT executive control score; CANT_PCR = the percentage of correct responses in the CANT; $CANT\_RT_{mdn}$ = the median response time in the CANT.

$^{\dagger}p < .10$. $^{*}p < .05$. $^{***}p < .001$.

3) Validity of the CANT

Concurrent validity of the CANT was examined to test the relationship between the scores of the CANT and other attention measures. In order to see how relevant the CANT is to the previously developed performance-based measure, correlation analysis was conducted between the CANT scores and FACT-Ⅱ-SR scores. Due to a missing data and an outlier in the FACT-Ⅱ-SR_CP data, data from 33 samples were used in the analysis. As Table 11 shows, CANT_O was weakly correlated with FACT-Ⅱ-SR_CP (r = .43, p = .01) and tended to show a positive correlation with FACT-Ⅱ-SR_CC (r = .32, p = .07). CANT_EC was weakly correlated with FACT-Ⅱ-SR_CC (r = .41, p = .02) and tended to show a positive correlation with FACT-Ⅱ-SR_CB (r = .33, p = .06). CANT_PCR showed a negative weak correlation with FACT-Ⅱ-SR_CP (r = -.36, p = .04). CANT_RT_{mdn} showed a moderate correlation with FACT-Ⅱ-SR_CC (r = .66, p = .00). On the other hand, the alerting score did not show significant correlation with any scores of FAST-Ⅱ-SR.

Concurrent validity of the CANT in relation to attention-related behavior ratings was examined. Pearson correlation analysis was conducted between the scores of the CANT and KADHDDS, and K-ADHDRS-T. Table 12 demonstrates the result of Pearson correlation analysis. CANT_O tended to show a positive correlation with the KADHDDS score (r = .32, p = .09). It indicated that students with a higher orienting score tended to show more inattentive behaviors at home. CANT_PCR showed a weak negative correlation with the KADHDDS score (r = -.40, p = .03). It indicated that students with higher CANT_PCR showed less inattentive behaviors at home. Except for this, none of the CANT scores showed significant correlations with the KADHDDS score.

Table 11

Result of Pearson Correlation Analysis between the Scores of CANT and FACT-Ⅱ-SR (n = 33)

	FACT-Ⅱ-SR_CP	FACT-Ⅱ-SR_CC	FACT-Ⅱ-SR_CB
CANT_A	.04	-.16	-.21
CANT_O	.43*	.32†	-.03
CANT_EC	-.03	.41*	.33†
CANT_PCR	-.36*	.20	.11
$CANT\_RT_{mdn}$	-.03	.66***	.22

Note. CANT_A = the CANT alerting score; CANT_O = the CANT orienting score; CANT_EC = the CANT executive control score; CANT_PCR = the percentage of correct responses in the CANT; $CANT\_RT_{mdn}$ = the median response time in the CANT; FACT-Ⅱ-SR_CP = the power of concentration performance in the FACT-Ⅱ-SR; FACT-Ⅱ-SR_CC = the concentration correctness in the FACT-Ⅱ-SR; FACT-Ⅱ-SR_CB = the concentration balance in the FACT-Ⅱ-SR.

$^{\dagger}p < .10$. $^{*}p < .05$. $^{***}p < .001$.

Table 12

Result of Pearson Correlation Analysis between the Scores of CANT, KADHDDS, and K-ADHDRS-T (n = 34)

	KADHDDS (n = 30)	K-ADHDRS-T (n = 34)
CANT_A	-.05	-.14
CANT_O	.32†	.01
CANT_EC	.04	.33†
CANT_PCR	-.40*	.09
$CANT\_RT_{mdn}$	.07	-.11

Note. CANT_A = the CANT alerting score; CANT_O = the CANT orienting score; CANT_EC = the CANT executive control score; CANT_PCR = the percentage of correct responses in the CANT; $CANT\_RT_{mdn}$ = the median response time in the CANT.

$^{\dagger}p < .10$. $*p < .05$.

Table 12 also shows the relationship between the CANT scores and the K-ADHDRS_T scores. Due to missing data, the data from 34 samples were used in the analysis. CANT_EC showed a weak positive correlation with the K-ADHDRS-T score (r = .33, p = .06). It indicated that students with a higher executive control score tended to show more inattentive behaviors at school. Other than this, no scores showed a meaningful relationship with the K-ADHDRS-T score.

4. Discussion

Study 1 explored the quality of the items, reliability, and validity of the CANT in Korean samples. The items were analysed in terms of their discrimination and validity. Reliability of the test was evaluated on internal consistency and stability over time. Validity was examined in terms of concurrent validity. The main findings of Study 1 are as follows.

Firstly, the results of item analysis indicate that the CANT items contribute the CANT discrimination. As the CANT was developed to test individual differences in attention, this result provides evidence for the quality of the CANT items. The validity of the CANT items appears appropriate. Conditions with no cue generate the longest response time, and conditions with spatial cues lead to the fastest reaction. Incongruent flankers distract subjects, thus slowing down the response speed. The double cue, center cue, and spatial cue change the mental state of the subjects into alerted so that as soon as the stimulus appears, they can respond to it as fast as possible. On the other hand, subjects do not maintain their vigilant state in the no cue condition, thus slowing down their response to a stimulus suddenly presented.

The spatial cue provides subjects information of where to orient their attention, thus overtly or covertly shifting their attention to the designated location efficiently. Whereas double cues and center cue do not help the orienting network operate efficiently, thus making attention wander around the fixation mark until the target appears. Incongruent flankers play a role in creating conflict. Subjects have to make a mental effort to resolve the conflict between the target and surroundings, which results in slower response speed. These results fit with the previous results. Rueda et al. (2004) also found a significant cue effect and flanker effect in response time. The accuracy data also reveals the effect of flanker conditions. Incongruent flankers decrease the percentage of correct responses. The result is also in line with Rueda et al. (2004)'s findings. Inconsistent flankers distract the subjects and lead them to make errors. These results demonstrate that the CANT items are valid in terms of their effects on manipulating the hypothesized condition with Korean samples.

Secondly, the reliability of the CANT ranges from insufficient to excellent, depending on the variables. The Cronbach's alpha for the accuracy and response time of the CANT was excellent. The split-half reliability of the median RT was excellent, and that of the percentage of correct response was good. However, alerting, orienting, and executive control show week correlations between the blocks. This result is partially consistent with the result in Rueda et al. (2004). Rueda et al. (2004) reported low to moderate split-half reliability for the three attention network scores and excellent split-half reliability estimates for the median response time and accuracy. The gap between the reliability estimates could be due to the different splitting methods they employed. While this study compared the scores between three blocks, Rueda et al. (2004) divided the first experimental block data into odd and even trials and calculated the split-half reliability coefficient. Nevertheless, the split reliability of the CANT scores is similar in two different cultures.

The test-retest reliability of the median RT is excellent, whereas those of alerting, orienting, executive control, and the percentage of correct response are not enough. This result is different from Rueda et al. (2004)'s study, in which the test-retest reliability appeared low. This could be due to the different time gaps between the first and second test in the studies. Rueda et al. (2004) retested 6.5 months later, whereas the test was conducted for the second time 10-14 days later in this study. However, the result of this study is in line with the ANT's test-retest reliability, which was reported as ranging from low to moderately high (Wang et al., 2015).

The low reliability of the CANT brings about some concerns since it weakens the statistical power of the results of studies where it could be applied. Some researchers have claimed that the low reliability of the ANT could be attributed to systematic error (MacLeod et al., 2010; Wang, et al., 2015). Reliability refers to the proportion of the total variance by the true variance. That is, the higher the reliability of a test, the greater the total variance caused by the true variance. However, a test's total variance is the sum of the true variance and error variance (Cohen et al., 2013). Therefore, if a test's reliability is low, it could be due to the error variance resulting from other error sources such as measurement, administration, scoring, and interpretation.

The design of the ANT has been attributed as a source of low reliability (Wang et al., 2015), and this argument can be applied to the CANT since its' design replicates the ANT's block design. As discussed, alerting, orienting, and executive control networks have their independent roles but work interactively in a task. Even though no inter-correlations among three networks and cue and flanker interactions were found in this study, inter-play of them has been found in some studies (Johnson et al., 2008; McConnell & Shore, 2011). The CANT trials randomly present one of the four cue types and one of the three flanker types. For example, incongruent flankers are presented with no cue, center cue, double cues, and spatial cues. The response time of incongruent flanker conditions is calculated across all the cue conditions, thus being influenced by other network components, i.e., alerting and orienting. This indicates that the executive control score is likely to be contaminated by other networks. Therefore, the influence of other networks could be one of the sources of systematic errors.

Additionally, attention network scores computed by subtraction could be another issue. The computation has strong face validity for achieving mathematical independence (MacLeod et al., 2010). However, considering that the overall median RT showed a consistent and stable result across the blocks of the test and two different times, the equation to obtain attention network scores based on RT might need an alternative. Wang et al. (2015) have suggested a revised equation isolating the attention system from the overall RT. It would be valuable to test the alternative equation or create a more reliable equation in future studies.

Lastly, the concurrent validity tests demonstrates the CANT's ecological validity. The CANT scores are correlated with the scores of another performance-based attention measure, FACT-Ⅱ-SR. The orienting score and percentage of correct responses are significantly correlated with the speed estimate on FACT-Ⅱ-SR. The executive control score and median RT have low to moderate correlation with the accuracy estimate on FACT-Ⅱ-SR. The executive control score is weakly correlated with the attentional balance-related score of the FACT-Ⅱ-SR. Even though FACT-Ⅱ-SR is not devised using the flanker paradigm and cue paradigm, the CANT appears to share a certain amount of variance of FACT-Ⅱ-SR. In the FACT-Ⅱ-SR, subjects are required to move to the next target after responding to the current target. This involves attention shift which includes disengaging from the current target, locating the next target, and shifting attention to it.

Therefore, students with a high orienting score could be disadvantaged in the FACT-Ⅱ-SR in terms of speed, thus resulting in a slower performance. Additionally, the speed-accuracy trade phenomenon seem to play a role in the result. The percentage of correct responses in the CANT is negatively correlated with the speed performance, i.e., FACT-Ⅱ-SR power of concentration performance, whereas the median response time in the CANT is positively correlated with the accuracy performance, i.e., FACT-Ⅱ-SR concentrations' correctness. This implies that the speed and the accuracy performance in the CANT predict the speed and the accuracy in the FACT-Ⅱ-SR.

Even though the correlation coefficients are statistically significant, they are still low to moderate estimates. That is, there are still variances of FACT-Ⅱ-SR that the CANT can not explain. The low to moderate correlation coefficients between the CANT and FACT-Ⅱ-SR could be due to the difficulty in operationally defining attention. As mentioned, attention is an umbrella term involving multiple aspects of mental processing. The CANT defines attention as attention networks involving achieving and maintaining alertness, selecting target information from sensory input, and resolving conflict among responses (Fan et al., 2002). On the other hand, FACT-Ⅱ-SR defines attention as speed, accuracy, the performance balance when discriminating targets and non-targets. Although both of them calculate their values on the basis of response time and accuracy, the way they operationally define attention differs. However, since both of them are direct measures to evaluate attention, there should be attention mechanisms commonly occurring. Hence, it is important to clarify the attention components of interest before selecting attention measures.

The CANT scores are meaningfully correlated with the attention-related behavior rating scores. The executive attention network score is positively correlated with the teachers' rating score. Although there is no study investigating criterion validity of the CANT, Beck et al. (2008) reported a moderate to high correlation between the ANT scores of brain-damaged patients and the attention-related questionnaire score rated by clinic staffs. However, Beck et al. (2008) did not use the subtraction formula suggested by Fan et al. (2002) but used the separate RT data in different cue and flanker condition. Even though the subjects in this study were typically-developing students, the executive control score of the CANT demonstrates its possible influence on students inattentive behaviors in school. This result implies that the executive control assessed in the CANT could account for the underpinning mechanism of inattentive behaviors of normal children. Laboratory tasks involving monitoring and resolving conflicts, such as stroop task and spatial conflict tasks have been used in studies on effortful action control (Simonds, Kieras, Rueda, & Rothbart, 2007). Attentional control is considered to be crucial in cognitive, emotional, and behavioral regulation (Luszczynska et al., 2004; Ochsner & Gross, 2005). The CANT asking students to decide the direction of the target fish surrounded by incompatible distracters could account for the behavioral dysregulation of inattentive students.

The orienting score is positively correlated with the parents' rating scores. The result implies that orienting deficit influences inattentive behaviors of students, at least at home. A high orienting score indicates difficulty in disengaging attention from the fixation mark. Adults with dyslexia also have been reported to have impaired attentional orienting (Buchholz & Davies, 2007). Disengagement difficulty has also been found in patients with parietal lesions (Fan & Posner, 2004). Therefore, it would be valuable to investigate how orienting affects academic achievement of children.

However, the alerting score did not show statistically significant relationship with the survey-based scores. This might be because the two types of attention assessments measure different types of alerting. The alerting score of the CANT represents phasic alertness, which is the rapid change in attention due to infrequent and unexpected events as warning signals followed by stimulus presentation. On the other hand, the items in the questionnaires measure tonic alertness, which refers to intrinsic arousal involved in sustaining attention for minutes to hours (DeGutis & Van Vleet, 2010). Tonic alertness rather than phasic alertness seems to affect inattentive behaviors.

Study 1 has several limitations. Firstly, it only dealt with a limited number of validity and reliability indices. It is important to test several aspects of psychometric features to assure a test's reliability and validity. This study, however, did not test other measures such as inter-scorer reliability and predictive validity, hence requiring more rigorous psychometric evaluation on the CANT in the future. Secondly, the sampling method and the sample size in this study restrict generalizing the results. The school was selected by convenient sampling, and the samples were all volunteers. Since it was required to obtain agreement from the participants and their legal guardians, voluntary sampling-a non-probability sampling method- was inevitable for this study. Additionally, only 71 students participated in this study. Performance-based tests take a long time and are hard to conduct simultaneously with a large number of subjects as survey-based tests can. However, to attain a more generalizable result, it is crucial to use as many samples as possible recruited by random sampling methods. Future studies should accommodate more samples to test other psychometric characteristics of the CANT.

This study, however, contributes to attention literature. Several performance-based attention measures have been developed and examined their validity. However, due to the lack of their operational definition of attention constructs in the studies, it was unclear which values to select to make inferences about attention. On the other hand, the attention network theory is a well-established and validated theory. Therefore, the ANT that measures the attention networks accordingly has been widely used in various academic fields. Although there were studies examining the reliability and validity of the measure, there were no studies on the CANT. Previous studies using the CANT failed to report its psychometric properties in their samples with diverse cultural backgrounds. Hence, this study provides a foundation for more use of the CANT in Korea by testing its psychometric quality and offering further suggestions to improve it.

Ⅳ. Study 2

1. Research Question

Study 2 aims to test the ecological validity of the CANT in an academic setting. The CANT is a test that is conducted in a controlled situation with possible contaminators eliminated. It uses pictures of fish in a clean blue background for students to respond to. In contrast, other sensory inputs, such as noise and other visual stimuli, distract students' attention in real classrooms. In order to demonstrate the ecological validity of the test, the performances in the test should be associated with those in real-world cognitive activity, in this study, mathematical word problem-solving. Based on the theoretical implications, the sub-questions for Study 2 were generated as follows:

Research Question: Does the CANT account for individual differences in the mathematical word problem performance?

Sub-questions:

1. Do the CANT performances significantly predict those on the mathematical word problem test after the controlling computation skill?

1-1. Does the CANT_PCR significantly predict the mathematical word problem-solving performances?

1-2. Does the $CNAT\_RT_{mdn}$ significantly predict the mathematical word problem-solving performances?

2. Does CANT_A significantly moderate the performances on the word problems?

2-1 Does CANT_A moderate the accuracy in word problem-solving?

2-1. Does CANT-A moderate the speed in word-problem solving?

3. Does CANT_O significantly moderate the performances on the word problems with irrelevant information?

3-1. Does the CANT_O moderate the accuracy in solving word problems with irrelevant information?

3-2. Does the CANT_O moderate the speed in solving word problems with irrelevant information?

4. Does the CANT_EC significantly moderate the performances on the word problems with inconsistent relational terms

4-1. Does the CANT_EC moderate the accuracy in solving word problems with inconsistent relational terms?

4-2. Does the CANT_EC moderate the speed in solving word problems with inconsistent relational terms?

5. Is the accuracy and speed pattern similar in the CANT and the word problem-solving?

2. Methodology

1) Participants

The target population for this study included fifth and sixth-grades students in a province, Korea. Participants were recruited in the same elementary school as in Study 1. Since there were multiple classes per grade, two classes in each grade were selected using the convenient sampling method, with the class examined in Study 1 excluded. The researcher visited the homerooms of the chosen classes and explained the purpose of the study, the tests the students would be asked to take, and rewards. Then permission letters were distributed to those who were willing to participate in the study. they were asked to return the letter with their parents or legal guardians' signature within a week. As a result, 89 letters were distributed, and 31 letters were returned. After excluding missing data of a student, 30 students' data were used in the analysis. The characteristics of the subjects are described in Table 13.

Table 13

Demographic Characteristics of Subjects

Grade	Number of Subjects(%)	Number of Male Subjects(%)	Number of Female Subjects(%)
5	19(61)	9(64)	10(59)
6	12(39)	5(36)	7(41)
Total	31(100)	14(100)	17(100)

2) Measurement

(1) The Child Attention Network Test (CANT)

The CANT was employed in the same way as descried as in Study 1.

(2) Mathematical Achievement Test (MAT)

The mathematical Achievement Test was composed of two parts: a math computation test (MCT) and a mathematical word problem test (MWPT). MWPT was composed of one-step word problems using arithmetic operations of addition, subtraction, multiplication, and division using integers. There were twenty word problems in total. All of them were extracted or modified from the previous studies on word problem solving (An, 2013; Jung, 2015; Lee, 2005; Sim, 2006). Five of them (items 5, 11, 14, 16, and 19) included irrelevant information for solving the problems, thus involving orienting manipulation (MWPT_O). Seven of them (items 1, 2, 7, 10, 17, 18, and 20) contained inconsistent relational terms to describe the unknown variables, thus involving executive control manipulation (MWPT_EC). Four of them (items 4, 9, 12, and 15) contained irrelevant information and inconsistent relational terms, thus involving orienting and executive control manipulation (MWPT_OEC). Four of them (items 3, 6, 8, and 13) were filler questions without manipulation (MWPT_NO). The MCT consisted of simple arithmetic questions. It was used to control the degree of automation in computation skills on the performance in the MWPT. There were ten questions of addition, subtraction, multiplication, and division. A practice problem was provided to explain how to register answers and proceed to the next question in each part. The outcome of the mathematical achievement test was based on the response time and number of correct responses.

3) Data Collection

The data were collected in September of 2020. Students were invited to a classroom with laptops set up on the desks. They took the CANT first, as described in Study 1. Then they took the MCT on the same laptop. They looked at the math problems presented on the screen, then entered their answers using the number pads on the keyboard. After registering their answers, the subjects clicked on the "다음문제" meaning "Next Question" button using the mouse. A piece of empty regular-sized paper and a pencil were provided in case the subjects would like to use for calculation. A practice quiz was given, followed by ten simple arithmetic questions. The researcher explained to the participants how to take the math tests individually. Finally, they took the MWPT on the same laptop in the same way as taking the MCT. It took less than 3 minutes for the MCT and approximately 10 minutes for the MWPT. The percentage of correct responses (PCR) and the mean response time for correct responses (RT_{mean}) were collected for later analysis. All the math questions were presented using the Inquisit Lab 5 software.

4) Data Analysis

Descriptive statistic analysis was conducted to calculate means, standard deviations, and correlation coefficients between the scores. In order to answer Sub-question 1, Pearson correlation analysis was conducted to determine the correlation between speed and accuracy performances on the CANT and MWPT. After detecting significant correlations, hierarchical linear regression analysis was conducted using the performances in MWPT as a dependent variable and the accuracy and the median response time from the CANT as independent variables. The MCP performances were entered in the first step of the equation in order to identify the independent ability of the CANT variables in accounting for the WMPT performance.

In order to answer Sub-questions 2, the participants were divided into two groups based on their CANT_A. The median of CANT_As was used as a cut-off for separating groups. Then the relationships between CANT_A and MWPT_PCR and $MWPT\_RT_{mean}$ in each group were examined by Pearson correlation analysis and regression analysis.

In order to answer Sub-questions 3, the participants were divided into two groups based on their CANT_O. The median of CANT_Os was used as a cut-off for separating groups. Then the relationships between CANT_O and the percentage of correct responses in solving word problems with irrelevant information (MWPT_PCR_O) and the mean response time for correct answers in solving word problems with irrelevant information ($MWPT\_RT_{mean}\_O$) were examined by Pearson correlation analysis and regression analysis.

In order to answer Sub-questions 4, the participants were divided into two groups based on their CANT_EC. The median of CANT_ECs was used as a cut-off for separating groups. Then the relationships between CANT_EC and the percentage of correct responses in solving word problems with inconsistent relational terms (MWPT_PCR_EC) and the mean response time for correct answers in solving word problems with inconsistent relational terms ($MWPT\_RT_{mean}\_EC$) were examined by Pearson correlation analysis and regression analysis.

The speed and accuracy performance patterns in the CANT and the MWPT were visualized in a quadrant graph to answer Sub-question 5. The means of speed and accuracy performance in each test were used as centers for the x-axis and the y-axis of the graph. The performances were marked on the quadrants and compared.

Correlation coefficients of less than .30 were interpreted as neglectable correlation, coefficients .from .30 to .49 were interpreted as weak correlation, coefficients from .50 to .69 were considered as moderate correlation, and coefficients of .70 and more were interpreted as high correlation (Mukaka, 2012). The statistical significance cut-off was .10 due to the small sample size in the study. SPSS was used for statistical analysis.

3. Result

1) Result of Descriptive Statistic Analysis

After excluding an outlier, data from 29 samples were analyzed. Table 14 presents the descriptive statistics for the performances on both measures. The attention network scores demonstrated correlations with the math achievement test results. CANT_A showed a moderate positive correlation with MWPT_OEC_PCR (r = .52, p = .01) and a weak correlation with MCT_RT$_{mean}$ (r = .42, p = .02). It tended to show a weak negative correlation with MWPT_OEC_PCR (r = −.31, p = .10) and a weak positive correlation with MWPT_NO_RT$_{mean}$ (r = .35, p = .06). CANT_O tended to show weak positive correlations with MWPT_O_PCR (r = .35, p = .06) and MWPT_O+OEC_PCR (r = .34, p = .06).

CANT_EC showed positive correlations with the response times in the math test and negative correaltions with the accuracy. CANT_EC showed weak positive correlations with MWPT_NO_RT$_{mean}$ (r = .37, p = .04) and MCT_RT$_{mean}$ (r = .39, p = .04). It tended to show positive correlations with MWPT_RT$_{mean}$ (r = .33, p = .08), MWPT_EC_RT$_{mean}$ (r = .36, p = .05), MWPT_EC+OEC_RT$_{mean}$ (r = .35, p = .06). Although it did not reach to the significance level, CANT_EC appeared to show a positive correlation with MWPT_O+OEC_RT$_{mean}$ (r = .31, p = .10). However, it showed a moderate negative correlation with MWPT_NO_PCR (r = −.52, p = .01) and tended to show a negative correlation with MWPT_OEC_PCR (r = −.38, p = .05).

CANT_PCR appeared to show positive correlations with the accuracy of the math test. It showed a weak positive correlation with MWPT_PCR (r = .44, p = .02) and MWPT_NO_PCR (r = .41, p = .03). Additionally, it tended to show a positive correlation with MWPT_EC+OEC_PCR (r = .44, p = .02). However, it appeared to show negative correlations with the response times on the math test. It showed negative correlations with MWPT_RT$_{mean}$ (r = −.47, p = .01), MWPT_EC+OEC_RT$_{mean}$ (r = −.44, p = .02), MWPT_NO_RT$_{mean}$ (r = −.47, p = .02) and MCT_RT$_{mean}$ (r = −.60, p = .01). Although it did not reach to the significance level, CANT_PCR showed a negative correlation with MWPT_O_RT$_{mean}$ (r = −.31, p = .10).

Table 14

Result of Pearson Correlation Analysis between the Scores of the CANT and MAT (N = 29)

Var.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	–																				
2	.17	–																			
3	.46**	.32†	–																		
4	.01	.37*	−.07	–																	
5	.04	.18	.05	−.09	–																
6	.08	.18	−.11	.18	.41*	–															
7	.42*	−.17	.39*	−.60**	.25	.15	–														
8	−.10	.17	−.23	.44*	.19	.43*	−.30	–													
9	.27	−.20	.33†	−.47*	.18	−.09	68***	−.36†	–												
10	−.12	.35†	.05	.24	.19	.28	−.38*	.58**	−.26	–											
11	.19	−.13	.25	−.31†	−.08	−.22	.41*	−.62***	.81***	−.32†	–										
12	.00	.10	−.17	.17	.00	−.09	−.14	.25	−.09	−.11	−.08	–									
13	.13	−.07	.36†	−.11	.12	.00	.18	−.07	.22	−.10	.00	.09	–								
14	−.31†	−.05	−.38*	.06	.20	.08	−.08	.21	.02	−.03	−.04	.50**	−.03	–							
15	.52**	−.05	.30	−.09	.21	.19	.38*	−.10	.45	−.12	.41*	.09	.32†	.07	–						
16	−.07	.34†	−.01	.37*	.17	.35†	−.38*	.81***	−.30	.84***	−.43*	.13	−.02	.18	−.07	–					
17	.24	−.15	.31†	−.29	.04	−.29	.37*	−.50**	.87***	−.29	.88***	−.03	.11	.00	.41*	−.33†	–				
18	−.07	.01	−.16	.33†	.20	.42*	−.07	.89***	−.18	.25	−.51**	.30	−.03	.22	−.01	.58**	−.34†	–			
19	.23	−.17	.35†	−.44*	.20	−.15	.59**	−.34†	.91***	−.23	.65***	−.12	.29	−.02	.28	−.28	.83***	−.14	–		
20	−.04	.16	−.52**	.41*	−.11	.04	−.46*	.47*	−.54**	.18	−.48**	.24	−.08	.23	−.19	.31	−.58**	.20	−.57**	–	
21	.35†	−.11	.37*	−.47*	.29	.08	.76***	−.32†	.85***	−.21	.66***	−.08	.34†	.06	.53**	−.25	.60**	−.21	.64***	−.34†	–
M	56.52	42.12	67.00	96.43	509.40	94.48	7074.34	75.69	19821.77	82.07	21141.61	70.81	25545.10	71.55	25495.27	77.78	23153.68	69.28	22580.33	85.34	14869.17
SD	39.66	35.99	28.46	2.70	53.43	8.27	2396.27	12.30	5317.12	17.19	8424.84	23.92	9353.76	22.88	9453.86	16.53	7419.83	16.91	5659.74	15.69	5513.79

Note. 1 = the CANT alerting score; 2 = the CANT orienting score; 3 = the CANT executive score; 4 = the percentage of correct responses in the CANT; 5 = The median response time for the correct answers in the CANT; 6 = the percentage of correct responses in the MCT; 7 = the mean response time in the MCT. 8 = the percentage of correct responses in the MWPT; 9 = the mean response time for correct answers in the MWPT; 10 = the percentage of correct answers in the MWPT including irrelevant information; 11 = the mean response time for correct answers in the MWPT including irrelevant information; 12 = the percentage of correct responses in the MWPT inclduing inconsistent relational terms; 13 = the mean response time in the MWPT including inconsistent relational terms; 14 = the percentage of correct responses in the MWPT including irrelevant information and inconsistent relational terms ; 15 = the mean response time for correct responses in the MWPT including irrelevant information and inconsistent relational terms; 16 = the percentage of correct responses in the MWPT with irrelevant information + with irrelevant information and inconsistent relational terms; 17 = the mean response time for correct answers in the MWPT with irrelevant information + with irrelevant information and inconsistent relational terms; 18 = the percentage of correct response in the MWPT with inconsistent relational terms + with irrelevant information and inconsistent relational terms; 19 = The mean response time in the MWPT with inconsistent relational terms + with irrelevant information and inconsistent relational terms; 20 = the percentage of correct responses in the MWPT_without manipulation; 21 = the mean response time for correct anwers in the MWPT without manipulation.

†p 〈 .10. *p 〈 .05. **p 〈 .01. ***p 〈 .001.

Lastly, $CANT\_RT_{mdn}$ showed a weak positive correlation with MCT_PCR (r = .41, p = .03). $CANT\_RT_{mdn}$ showed neglectable positive correlations with the other scores except $MWPT\_O\_RT_{mean}$ (r = −.08, p = .69) and MWPT_NO_PCR (r = −.11, p = .59).

2) The Influence of the CANT Accuracy and Speed on the MWPT Performances

In order to test if the accuracy of the CANT can predict the accuracy and speed in the MWPT, hierarchical regression analysis was conducted. The accuracy and speed performances of the MCT were used as confounding variables to be controlled.

As Table 14 indicates, the CANT_PCR showed a significant correlation with the MWPT's accuracy (r = .44, p = .02). The MCT_PCR also showed a positive correlation with MWPT_PCR (r = .42, p = .03). These results demonstrate that the CANT's accuracy could account for a unique variance of the MWPT's accuracy. In order to test if the accuracy performance of the CANT can predict that on the MWPT, hierarchical regression analysis was conducted.

As Table 15 summarizes, CANT_PCR had a positive correlation with the number of the correct answer in the MWPT (β = .37, p = .03). The first model indicates MCT_PCR accounted for 17% of the variance of MWPT_PCR. The second model demonstrates MCT_PCR and CANT_PCR explained 31% of the accuracy performance in the MWPT. CANT_PCR appeared to explain 14% of MWPT_PCR accuracy after controlling the effect of MCT_PCR. These results demonstrate that the CANT accuracy accounts for a unique variance of the MWPTs' accuracy.

CANT_PCR was negatively correlated with the $MWPT\_RT_{mean}$ (r = −.47, p = .01). In order to test if it accounts for a unique variance of $MWPT\_RT_{mean}$, hierarchical analysis was conducted entering MCT_RT_{mean} as a confounding variable. MCT_RT_{mean} accounted for 46% of the variance of $MWPT\_RT_{mean}$, $\beta = .69$, $F(1, 27) = 23.38$, $p < .001$. After controlling the MCT_RT_{mean}, CANT_PCR appeared to account for 1% of the $MWPT\_RT_{mean}$ variance. However, CANT_PCR showed a moderate negative correlation with MCT_RT_{mean} ($r = -.60$, $p = .01$), another regression test was conducted in order to see if MCT_RT_{mean} mediates the relationship between CANT_PCR and $MWPT\_RT_{mean}$. As Table 16 presents, CANT_PCR accounted for 22% of the $MWPT\_RT_{mean}$, $\beta = -.47$, $F(1, 27) = 7.71$, $p = .01$, when it was the only variable entered in the regression. However, after entering MCT_RT_{mean} in the regression, its influence decreased, $\beta = -.10$, $F(1, 27) = 23.38$, $p = .60$. This proved the mediating role of MCT_RT_{mean} in the relationship between CANT_PCR and $MWPT\_RT_{mean}$.

$CANT\_RT_{mdn}$ showed a positive correlation with MWPT_PCR ($r = .19$, $p = .32$) and MCT_PCR ($r = .41$, $p = .03$). Since MCT_PCR were positively correlated with MWPT_PCR ($r = .43$, $p = .02$), regression analysis was

Table 15

Result of Hierarchical Regression Analysis for CANT_PCR Predicting MWPT_PCR (N = 29)

Variable	Model 1			Model 2		
	B	β	*SE*	*B*	β	*SE*
Constant	16.04		24.56	−138.92		72.20
MCP_PCR	.63	.43*	.26	.53	.36*	.25
CANT_PCR	–		–	1.70	.37*	.75
R^2	.18*				.32*	
ΔR^2					.14*	

Note. MCP_PCR = the percentage of correct responses in the MCP; CANT_PCR = the percentage of correct responses in the CANT.
$^{*}p < .05$.

conducted to see if MCT_PCR mediates the relationship between $CANT\_RT_{mdn}$ and MWPT_PCR. However, the regression model did not appear significant, $F(1, 27) = 1.05$, $p = .31$, therefore no further analysis was followed.

Table 16

Result of Hierarchical Regression Analysis for MCT_RT_{mean} Mediating CANT_PCR and $MWPT\_RT_{mean}$ (N = 29)

Variable	Model 1			Model 2		
	B	β	*SE*	*B*	β	*SE*
Constant	109461.43		32302.84	28167.59		35810.02
CANT_PCR	-929.569	-.47*	334.86	-188.08	-.10	352.99
MCT_RT_{mean}				1.38	.62**	.40
R^2	.22*			.47***		
ΔR^2				..22		

Note. CANT_PCR = the percentage of correct responses in the CANT ; MCT_RT_{mean} = the mean response time for correct responses in the MCT.

*$p < .05$. **$p < .01$. ***$p < .001$.

Table 17

Result of Regression Analysis for $CANT\_RT_{mdn}$ Predicting $MWPT\_RT_{mean}$ (N = 29)

	B	*SE*	β	*F*	*p*
Constant	10830.56	2330.92		.88	.36
$CANT\_RT_{mdn}$	17.651	.31	.18		

Note. $CANT\_RT_{mdn}$ = the median response time for correct responses in the CANT; $MWPT\_RT_{mean}$ = the mean response time for correct responses in the MWPT ; MCT_RT_{mean} = the mean response time for correct responses in the MCT.

$CANT\_RT_{mdn}$ showed a positive correlation with $MWPT\_RT_{mean}$ (r = .18, p = .35) and MCT_RT_{mean} (r = .25, p = .19). As Table 17 shows, the result of regression analysis demonstrated that the regression of $CANT\_RT_{mdn}$ accounting for $MWPT\_RT_{mean}$ was not statistically significant, F(1, 27) = .88, p = .36. Therefore, no further analysis was followed.

3) Moderating effect of CANT_A on the MWPT Performances

In order to see the moderating effect of CANT_A on the MWPT performance, the participants were divided into two groups; the high alerting group and the low alerting group. The correlation analysis between CANT_A and MWPT_PCR and $MWPT\_RT_{mean}$ were conducted on each group. As Table 18 shows, CANT_A showed a positive correlation with MWPT_PCR (r = .07, p = .82) in the low CANT_A group, and a positive correlation (r = .20, p = .48) in the high CANT_A group. The regression analysis of CANT_A predicting MWPT_PCR did not appear statistically significant, $F(1, 13)$ = .06, p = .82 and $F(1, 14)$ = .54, p = .48, respectively. In addition, CANT_A showed neglectable correlations with $MWPT\_RT_{mean}$ in both groups (r = .05, p = .87 and r = −.07, p = .80, respectively). The regression of CANT_A predicting $MWPT\_RT_{mean}$ was not significant, $F(1, 13)$ = .03, p = .87 and $F(1, 14)$ = .07, p = .80, respectively.

Table 18

Result of Pearson Correlation Analysis between CANT_A and the MWPT Performances in the High CANT_A and Low CANT_A Groups (N = 29)

	Low CANT_A Group	High CANT_A Group
	CANT_A	CANT_A
MWPT_PCR	.07	.20
$MWPT\_RT_{mean}$	.05	−.07

Note. CANT_A = the CANT alerting score; MWPT_PCR = the percentage of correct responses in the MWPT; $MWPT\_RT_{mean}$ = the mean response time for correct responses in the MWPT.

4) The Moderating Effect of CANT_O on Solving Word Problems with Irrelevant Information

Table 19 shows the result of Pearson correlation analysis between CANT_O and the accuracy and speed in solving word problems with irrelevant information. CANT_O showed positive correlations with MWPT_O_PCR (r = .28, p = .34) in the low CANT_O group and the high CANT_O group (r = .22, p = .42). It showed positive correlations with $MWPT\_O\_RT_{mean}$ in both groups (r = .14, p = .63 and r = .16, p = .57, respectively). Regression analysis were conducted in order to see if the influence of CANT_O on the MWPT_O performances varies among different CANT_O groups. However the regressions of CANT_O predicting the MWPT_O_PCR did not appear significant in both groups, $F(1, 13) = .99$, p = .34 and $F(1, 14) = .49$, p = .42, respectively. The regressions of CANT_O predicting the $MWPT\_O\_RT_{mean}$ also appeared not significant, $F(1, 13) = .24$, p = .63 and $F(1, 14) = .33$, p = .57, respectively.

Table 19

Result of Pearson Correlation Analysis between CANT_O and the MWPT Performances in the High CANT_O and Low CANT_O Groups (N = 29)

	Low CANT_O Group	High CANT_O Group
	CANT_O	CANT_O
MWPT_O_PCR	.28	.22
$MWPT\_O\_RT_{mean}$	.14	.16

Note. CANT_O = the CANT orienting score; MWPT_O_PCR = the percentage of correct responses in the MWPT with irrelevant information; $MWPT\_O\_RT_{mean}$ = the mean response time for correct responses in the MWPT with irrelevant information.

5) The Moderating Effect of CANT_EC on Solving Word Problems Including Inconsistent Relational Terms

In order to test if the influence on CANT_EC on the performance in solving word problems with inconsistent relational terms varies among groups with different CANT_EC, Pearson correlation analysis and regression analysis were employed. As Table 20 indicates CANT_EC showed a negative correlation with MWPT_EC_PCR in the low CANT_EC group (r = −.35, p = .22), but did not any correlation in the high CANT_EC group (r = .00, p = 1.00). It showed a negative correlation with MWPT_EC_RT$_{mean}$ in the low CANT_EC group (r = −.42, p = .14), but a positive correlation in the high CANT_EC group (r = .10, p = .73). Regression analysis was conducted to test if the impact of CANT_EC on the performance in the MWPT with inconsistent relational terms varies among students with different CANT_EC. However, the regression of CANT_EC predicting MWPT_EC_PCR appeared statistically insignificant in the low CANT_EC group, $F(1, 13)$ = 4.68, p = .22, and in the high CANT_EC group, $F(1, 14)$ = .00, p = 1.00. The regression of CANT_EC predicting MWPT_EC_RT$_{mean}$ also did not appear significant in the low CANT_EC group, $F(1, 13)$ = 2.57, p = .14, and in the high CANT_EC group, $F(1, 14)$ = .13, p = .73.

Table 20

Result of Pearson Correlation Analysis between CANT_EC and the MWPT Performances in the High CANT_EC and Low CANT_EC Groups (N = 29)

	Low CANT_EC Group	High CANT_EC Group
	CANT_EC	CANT_EC
MWPT_EC_PCR	−.35	.00
MWPT_EC_RT$_{mean}$	−.42	.10

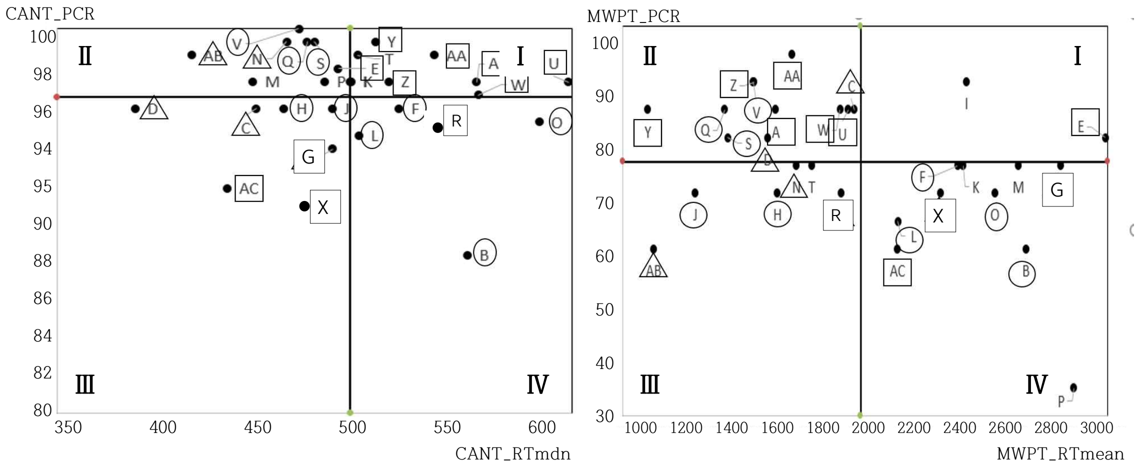
Note. CANT_EC = the CANT executive control score; MWPT_EC_PCR = the percentage of correct responses in the MWPT with inconsistent relational terms; MWPT_EC_RT$_{mean}$ = the mean response time for correct responses in the MWPT with inconsistent relational terms.

6) The Performance Pattern in the CANT and MWPT

The accuracy and speed performances of the CANT and MWPT are depicted in Figure 6. Three students (S, V, Q) in Quadrant Ⅱ showed the same pattern (high accuracy and fast speed), two students (J, H) in Quadrant Ⅲ showed the same pattern (low accuracy and fast speed), and four students (B, F, L, O) in Quadrant Ⅳ showed the same pattern (low accuracy and slow speed) in both tests. Eleven students (A, E, G, R, U, W, X, Y, Z, AA, and AC) showed the same accuracy pattern in the tests, and four students (C, D, N, and AB) showed the same speed pattern in the tests. Four students (I, K, M, and P) showed the opposite pattern (high speed and low accuracy in the CANT, but low speed and high accuracy in the MWPT). It can be inferred that most of the cases show a similar pattern in terms of either speed or accuracy performance in the two tests.

Figure 6

The Performance Pattern in the CANT and the WMPT



Note. CANT_PCR = the percentage of correct responses in the CANT; $CANT\_RT_{mdn}$ = the median response time in the CANT; MWPT_PCR = the percentage of correct responses in the MWPT; $MWPT\_RT_{mean}$ = the mean response time for correct responses in the MWPT. The alphabet letters indicate each sample. The samples in the circle showed the same speed and accuracy pattern in both tests. The samples in the rectangle showed the same accuracy pattern in both tests. The samples in the triangle showed the same speed pattern in both tests.

4. Discussion

Study 2 tested the ecological validity of the CANT in the context of mathematical word problem-solving. Based on the attention network theory and Mayer's (1980) model of mathematical problem-solving, it was assumed that the performance in the CANT would predict the performance in the word problem-solving. The main findings from Study 2 are as follows.

Firstly, the CANT accuracy performance significantly predicts the accuracy and response time in the word problem solving. The result indicates that the accuracy performance in the CANT can be generalized in the mathematical problem-solving situation. Studies using the ANT as an attention measure predicting mathematical achievement employed only the three attention network scores (Antonini et al., 2016; Gold et al., 2013), resulting in no correlations between the networks and math achievement. Furthermore, no studies on the relationship between mathematical achievement and attention measured by the CANT could be traced. Therefore, this result is valuable since it demonstrates the ecological validity of the CANT.

The finding also leads researchers to reconsider what the accuracy data indicates in terms of the attention networks' efficiency. The CANT's accuracy data was often neglected because the ANT was developed to mainly measure the three networks' efficiency. In some studies, the accuracy data did not show significant group differences (Redick & Engle, 2006), whereas other studies revealed individual differences in accuracy among participants of different ages and genders (Rueda et al., 2004; Westlye, Grydeland, Walhovd, & Fjell, 2010). Besides, Johnson et al. (2008) have found out that ADHD students showed a higher number of incorrect responses in incongruent conditions than typically developing peers, hence implying the accuracy performance could be used as an indicator for ADHD. Fan and Posner (2004), who participated in developing the ANT and CANT, suggested to consider the accuracy data in interpreting the efficiency of attention networks. Therefore, future studies need to find out further possibility for the CANT's accuracy performance to be generalized in other everyday life settings such as other school subjects and social activities.

On the other hand, the median RT of the CANT does not predict the accuracy and speed performances on mathematical word problems. Since no studies have investigated the relationship between the CANT's speed measure and mathematical performance, the explanation of this result could only be inferred from studies with another variable, processing speed, which seemingly shares a similar mental aspect with the response time. Processing speed can be defined as the efficiency or the speed at which one executes simple cognitive tasks (Passolunghi and Lanfranchi, 2012). The effect of processing speed on math achievement is inconsistent across studies. Bull and Johnston (1997) found that processing speed measured by a visual number matching, cross-out task, and perceptual-motor speed was the best predictor of arithmetic ability of 7-year-old children. Processing speed measured by a visual matching task predicted the outcome of math fact fluency test with answers ranging from o to 12 but did not predict the arithmetic word problem performance (Fuchs et al., 2006). Besides, processing speed measured by target reaction tasks did not predict the math performance of middle school students (Cai, Li, & Deng, 2013). Consequently, the influence of speed on the CANT could differ depending on subjects and math problems, suggesting further research on its role in diverse contexts.

Secondly, the scores of the CANT alerting, orienting, and executive control demonstrate different levels of ecological validity. The alerting network has a weak positive correlation with the speed performance in the word problems without manipulation and in the mathematical computation test. That is, students with a higher alerting score take a longer time to solve word problems without manipulation and in the simple mathematical computation problems. The result also demonstrates that alerting does not have meaningful correlation with the accuracy and speed performances in the mathematical word problem-solving. This indicates that the alerting score can be generalized mostly in simple mathematical problem-solving that requires no additional thinking strategy.

Thirdly, the orienting score shows a positive correlation with the accuracy in the mathematical word problems with irrelevant information. This indicates that students with a higher orienting score can be expected to make fewer errors in solving mathematical word problems with irrelevant information. Thus, the orienting score can be generalized in solving word problems with irrelevant information in terms of accuracy. However, although a meaningful correlation between the orienting score and the word problem-solving performances can be found, the orienting score's influence on solving word problems including irrelevant information is not different among groups with the different orienting scores. Therefore, it could be concluded that the generalizability of the CANT orienting score on the mathematical problem-solving is partially supported.

Fourthly, the executive control score is positively correlated with the speed performance in word problems with inconsistent relational terms. The result indicates that students who struggle in resolving conflicts in the CANT need more time to monitor inconsistent relational terms and to resolve the conflicts between the language and the operation required in the problem-solving. In addition, the executive control score is negatively correlated with the accuracy and positively correlated with speed in the word problems without manipulation. This indicates that the students who spend longer time monitoring and resolving conflicts in the CANT take a longer time and make errors in solving simple word problems. However, although a meaningful correlation between the CANT executive control score and the word problem-solving performances can be found, the CANT executive control score does not significantly predict the performance in solving word problems with irrelevant relational terms. Therefore, it could be concluded that the generalizability of the CANT executive control performance is partially supported.

However, the small size of correlation coefficients between the attention network scores and the performances in the word problems with manipulation needs to be noted. The alerting network score is not meaningfully correlated with the performance in the word problems with manipulation. The orienting score is only weakly correlated with the accuracy and shows a neglectable correlation with the response time for the word problems with irrelevant information. The executive control score is weakly correlated with the response time and shows a neglectable correlation with the response time in the word problems with inconsistent relational terms. One explanation for these results could be that there are other cognitive or motivational factors playing more crucial roles in solving complex word problems. Complexity can be determined by the number of steps the problem solver needs to take (Musso, Boekaerts, Segers, & Cascallar, 2019) and by the quality of the information provided in the test (Kim & Kim, 2006). Musso et al. (2019) have found out that the accuracy performance on high complexity/low difficulty items was not predicted by executive attention but was predicted by working memory and subjective competence. This indicates that there could be other cognitive and motivational factors than attention contributing to a complex math problem -solving. However, as this study demonstrates, alerting and executive control networks are significantly correlated with the performance in simple math problems: simple arithmetic computations and word problems without manipulation. In other words, the degree of attention's influence varies depending on the amount of cognitive demand that mathematical word problems require.

On the other hand, the small sample size in this study might have contributed to underestimating the correlations. Statistical power depends on the sample size, and the true population value could be under- or over-estimated by the sample size (Wolf, Harrington, Clark, & Miller, 2013). Antonini et al. (2016) found out the orienting score, along with alerting, was significantly correlated with the math achievement of 147 students at the age of 7 to 11. Hence future researchers would want to recruit more samples representing the target population to expand the knowledge this study has gained.

Lastly, the performance in the CANT demonstrates a similar pattern in the mathematical word problem-solving. Except for four cases, all the samples exhibit the same accuracy or speed pattern, sometimes both as expected. This indicates that there is a similar individual characteristic operating in the two cognitive tasks. Kim and Youn (2013) grouped young children's cognitive styles according to the result of the Matching Familiar Figure Test. As they proposed, the efficient group (high speed and high accuracy), impulsive group (high speed and low accuracy), reflective group (slow speed and high accuracy), and inefficient group (slow speed and low accuracy) were identified in this study. Furthermore, their cognitive styles detected in the CANT are similarly reflected on the MWPT. Therefore, the speed × accuracy pattern in the CANT can represent the pattern in mathematical problem-solving, thus demonstrating the ecological validity of the CANT.

There are several limitations in generalizing the findings in this study. Due to the small sample size and the items in the test, one participant or an item could considerably change the results. Therefore, future researchers would want to sample more students and generate more items measuring academic achievement-related variables. Confounding variables influencing math performance could be attributed to the limitations. In this study, the degree of computation automation was the only controlled confounding variable, reducing the internal validity of the test. There could be more variables influencing math achievements, such as IQ, emotional states, and motivation factors. Additionally, during the computerized MWPT test, more unexpected confounding variables were detected, such as the speed of typing number pads and the speed of locating the mouse pointer at the "Next" button on the screen. Also, the length of each word problem and the number of digits used in the problems should be controlled. Future studies need to control as many variables as possible to get a more precise contribution of attention to academic achievement.

Nonetheless, this study contributes to attention literature in several ways. First of all, it attempts to test the ecological validity of the CANT in the academic setting. Few studies have examined the relationship between the performance of the CANT and other academic achievement or school adjustment variables. By investigating the CANT as a predictor for mathematical achievement, this study suggests an alternative way to validate behavioral and neuropsychological tests. Secondly, this study provides insight into the mechanism of attention operation in the learning process. Even though the manipulated word problems revealed partial information about attention networks' operation, the performance for word problems with and without manipulation could be predicted by the attention network scores. This implication of the study provides a need for more in-depth investigation in the education field, thus expanding the scope of attention literature. This study also offers a useful tool that schools can accommodate for testing students' attention in relation to their academic achievement.

V. Conclusion

This research aims to evaluate the ecological validity of the CANT. Based on the results of item analysis, validity analysis, and reliability analysis with Korean samples, it can be concluded that the CANT consists of valid items that discriminate individual differences. The reliability of the CANT ranges from poor to excellent depending on the scores. In addition, the CANT scores showed meaningful correlations with the scores from another performance-based attention measure and behavioral rating scales. These results indicate that the CANT is a valid attention measure with quality items, but its reliability needs improvement. This research also examines the ecological validity of the CANT in an educational setting. By exploring the relationship between the CANT scores and the mathematical word problem-solving performances, this research reveals that the CANT result predicts the mathematical word problem-solving performance, thus providing evidence for generalizing the test result to academic achievement.

This research clearly illustrates the psychometric properties of the CANT with Korean samples, employing various indices. Even though the ANT is a widely used attention measure, the validity and reliability of its child version - the CANT- has not been thoroughly investigated. It is crucial to use a valid and reliable measure in studies to assure veracity and consistency of the knowledge they are seeking. Therefore, it is a valuable attempt to investigate the quality of the CANT utilizing multiple methods for further attention research with children in Korea. Furthermore, this research demonstrates the predictive power of the CANT on academic achievement. This research, thus, contributes to expanding the scope of the CANT's utility from a country and experimental neuropsychological settings to places with different cultural backgrounds and educational settings.

However, this research raises some questions. First of all, many attention assessments might not be measuring the same constructs. This could be why some of the CANT scores are not correlated with the those of the criterion measures in this study. As attention is an overarching term, including multiple aspects of the mental process, attention measures should clearly introduce which components of attention they are devised to assess. The meanings of the estimates that many attention measures yield need to be interpreted in relation to the definition of attention subconstructs.

Secondly, the interplay of three attention networks could be a critical phenomenon to be investigated. As attention network theory suggests, the three networks carry out specific functions, but oftentimes they operate interactively. However, the CANT randomly presents cueing and flanking conditions in the experimental blocks resulting in the unintended interplay among the networks. This could be the reason for the unexplained variance between the test blocks and tests at different two times. Therefore, a new block design that can fractionate three networks could improve the quality of the CANT.

Based on these conclusions, attention researchers can use the CANT in their studies investigating children's attention in Korea. It is a performance-based measure minimizing the influence of the examiner's bias on the result. It is a computerized test so that a large number of children can take the test simultaneously as long as the test environment is equipped. Furthermore, developing the CANT operating in the mobile environment would increase the accessibility of the tool. However, researchers must consider their operational definitions of attention or its subconstructs to ensure the appropriate use of the CANT.

Educators may use the CANT in school settings. Attention deficits have become a serious problem in schools. In order to obtain an objective estimate, educators can use the CANT to evaluate students' attention. Its easiness in administration and its automatic scoring system could also benefit the interested educators. Especially if the CANT's predictive validity is tested, it could be a useful tool in schools for screening and testing the effects of interventions for students. Additionally, creating more child-friendly stimulus or providing options to choose various stimulus for the test would motivate students to participate in the test.

Building upon this study, future studies could validate the CANT in different settings. The validation process does not finish at the test development stage but continues by further validating efforts or studies using the test of interest in relation to other variables. In order to ensure the CANT as an ecologically sound measure, its item quality, reliability, and validity should be further investigated in various cultures. Moreover, the CANT scores representativeness should be examined in various settings.

Furthermore, future researchers need to put effort into improving its reliability. As suggested above, the in-block design could be revised to eliminate noise. Standardizing test administration and different scoring methods could be utilized. Multiple methods can be employed to reduce the systematic error variance. This process would contribute to making the CANT more sound.

This research builds a foundation for future attention research with children. Examining psychometric properties is a prerequisite process before using assessment tools in studies or educational settings. By providing evidence for item appropriateness, reliability, and validity of the CANT, this research allows future attention researchers to have clear ideas of when and how to accommodate the tool for their purpose. Considering the ANT's wide application in diverse academic fields, the CANT would also further expand attention literature.

Additionally, this research emphasizes the importance of defining attention or attention subconstructs in studies. Estimates from various performance-based attention measures have been used in attention studies. However, the latent variables that the values represent were not explicitly stated. On the contrary, the CANT was developed based on the attention network theory, and its estimates are operationally well-defined and tested in previous studies. The CANT would be a good example of attention assessment showing a clear idea of which aspects of attention it attempts to measure.

Lastly, this research provides insights into the mechanism of attention operation in mathematical problem-solving. This implies that not only mathematical achievement but other factors influencing successful school adjustment could be subject to explore in relation to attention. It would be possible to explain various problems of students and to develop intervention programs employing the CANT.

References

Adólfsdóttir, S., Sørensen, L., & Lundervold, A. J. (2008). The attention network test: A characteristic pattern of deficits in children with ADHD. *Behavioral and Brain Functions, 4*(1), 1-9.

Allport, A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. *Perspectives on Perception and Action, 15*, 395-419.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standard for educational and psychological testing.* NW: American Educational Research Association.

Anderson, B. (2011). There is no such thing as attention. *Frontiers in Psychology, 2*, 1-8.

Anobile, G., Stievano, P., & Burr, D. C. (2013). Visual sustained attention and numerosity sensitivity correlate with math achievement in children. *Journal of Experimental Child Psychology, 116*(2), 380-391.

An, S. K. (2013). *The effects of instructional models, visual and spatial working memory, and representation aids on students' math word problem solving*. Doctoral dissertation, Chonnam National University.

Ansari, D., Lyons, I. M., van Eimeren, L., & Xu, F. (2007). Linking visual attention and number processing in the brain: The role of the temporo-parietal junction in small and large symbolic and nonsymbolic number comparison. *Journal of Cognitive Neuroscience, 19*(11), 1845-1853.

Antonini, T. N., Kingery, K. M., Narad, M. E., Langberg, J. M., Tamm, L., & Epstein, J. N. (2016). Neurocognitive and behavioral predictors of math performance in children with and without ADHD. *Journal of Attention*

Disorders, 20(2), 108-118.

Ashkenazi, S., & Silverman, S. (2017). Multiple skills underlie arithmetic performance: A large-scale structural equation modeling analysis. *Journal of Numerical Cognition, 3*(2), 496-515.

Atkinson, J., Braddick, O., & Breckenridge, K. (2010). Components of attention in normal and atypical development. *Journal of Vision, 10*(7), 492-492.

Bae, J. W., Cheung, S. D., Lee, J. B., Park, H. B., Kim, J. S., & Song, C. J. (1997). The study of reliability and validity of Attention Deficit Disorder Evaluation Scale-School Version(ADDES-SV). *Yeungnam University Jounal of Medicine, 14*(2), 415-429.

Beck, L., Heusinger, A., Boecker, M., Niemann, H., & Gauggel, S. (2008). Convergent and predictive validity of two computerized attention tests in brain-damaged patients. *Zeitschrift für Neuropsychologie, 19*(4), 213-222.

Berger, A., & Posner, M. I. (2000). Pathologies of brain attentional networks. *Neuroscience and Biobehavioral Reviews, 24*, 3-5.

Broadbent, D. E. (1958). *Perception and communication*. London: Oxford University Press.

Brunning, R. H., Schraw, G. J., & Norby, M. M. (2011). *Cognitive psychology and instruction*. London: Pearson.

Buchholz, J., & Davies, A. A. (2008). Adults with dyslexia demonstrate attentional orienting deficits. *Dyslexia, 14*(4), 247-270.

Bull, R., & Johnston, R. S. (1997). Children's arithmetical difficulties: Contributions from processing speed, item identification, and short-term memory. *Journal of Experimental Child Psychology, 65*(1), 1-24.

Cai, D., Li, Q. W., & Deng, C. P. (2013). Cognitive processing characteristics of 6th to 8th grade Chinese students with mathematics learning disability: Relationships among working memory, PASS processes, and processing speed. *Learning and Individual Differences, 27*, 120-127.

Callejas, A., Lupianez, J., Funes, M. J., & Tudela, P. (2005). Modulations among the alerting, orienting and executive control networks. *Experimental Brain Research, 167*(1), 27-37.

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review, 13*(4), 181-197.

Choi, B. M., Kim, W. S., & Park, N. S. (1994). Use of T.O.V.A. in assessing attentional deficit of ADHD children. *Journal of Korean Neuropsychiatric Association, 33*(6), 1306-1320.

Cohen, R. J., Swerdlik, M. E., & Sturman, E. D. (2013). *Psychological testing and assessment* (8th ed.) New York: McGraw-Hill.

Daroczy, G., Wolska, M., Meurers, W. D., & Nuerk, H. C. (2015). Word problems: A review of linguistic and numerical factors contributing to their difficulty. *Frontiers in Psychology, 6*, 1-13.

Das, J. P. (2002). Attention: Concept, tests and teacher ratings. *Psychology and Developing Societies, 14*(2), 241-260.

DeGutis, J. M., & Van Vleet, T. (2010). Tonic and phasic alertness training: A novel behavioral therapy to improve spatial and non-spatial attention in patients with hemispatial neglect. *Frontiers in Human Neuroscience, 4*, 1-17.

Demetriou, A., Kui, Z. X., Spanoudis, G., Christou, C., Kyriakides, L., & Platsidou, M. (2005). The architecture, dynamics, and development of mental processing: Greek, Chinese, or Universal?. *Intelligence, 33*(2), 109-141.

DuPaul, G. J., Power, T. J., Anastopoulos, A. D., Reid, R., McGoey, K. E., & Ikeda, M. J. (1997). Teacher ratings of attention deficit hyperactivity disorder symptoms: Factor structure and normative data. *Psychological Assessment, 9*(4), 436-444.

Fan, J., Gu, X., Guise, K. G., Liu, X., Fossella, J., Wang, H., & Posner, M. I. (2009). Testing the behavioral interaction and integration of attentional networks. *Brain and Cognition, 70*(2), 209-220.

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *Neuroimage, 26*(2), 471-479.

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience, 14*(3), 340-347.

Fan, J., & Posner, M. (2004). Human attentional networks. *Psychiatrische Praxis, 31*(2), 210-214.

Fossella, J., Sommer, T., Fan, J., Wu, Y., Swanson, J. M., Pfaff, D. W., & Posner, M. I. (2002). Assessing the molecular genetics of attention networks. *BMC Neuroscience, 3*(1), 1-11.

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., ... & Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology, 98*(1), 29-43.

Gabrieli, J. D. E., & Whitfield-Gabrieli, S. (2007). Attention to neglect. *Neuron, 53*(6), 776-777.

Gold, A. B., Ewing-Cobbs, L., Cirino, P., Fuchs, L. S., Stuebing, K. K., & Fletcher, J. M. (2013). Cognitive and behavioral attention in children with math difficulties. *Child Neuropsychology, 19*(4), 420-437.

Hancock, P. A., Oron-Gilad, T., & Szalma, J. L. (2007). Elaborations of the multiple-resource theory of attention. In A. F. Kramer, D. A. Wiegmann, & A. Kirlik (Eds.), *Attention: From theory to practice* (pp. 45-56). New York: Oxford University Press.

Hart, S. A., Petrill, S. A., Willcutt, E., Thompson, L. A., Schatschneider, C., Deater-Deckard, K., & Cutting, L. E. (2010). Exploring how symptoms of attention-deficit/hyperactivity disorder are related to reading and mathematics performance: General genes, general environments.

Psychological Science, 21(11), 1708-1715.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New York: John Wiley.

Hegarty, M., Mayer, R. E., & Green, C. E. (1992). Comprehension of arithmetic word problems: Evidence from students' eye fixations. *Journal of Educational Psychology, 84*(1), 76-84.

Ishigami, Y., & Klein, R. M. (2009). Are individual differences in absentmindness correlated with individual differences in attention?. *Journal of Individual Differences, 30*(4), 220-237.

Jaekel, J., Wolke, D., & Bartmann, P. (2013). Poor attention rather than hyperactivity/impulsivity predicts academic achievement in very preterm and full-term adolescents. *Psychological Medicine, 43*(1), 183-196.

James, W. (1890). *The principles of psychology*. New York: Henry Holt and Company.

Jeong, G. M., & Song, J. H. (2019). Meta-analysis of the reliability and validity of studies on development of attention measures in Korea. *The Korean Journal of Educational Psychology, 33*(2), 171-191.

Jeong, G. M., Song, J. H., & Kim, H. S. (2019). An analysis of a keyword network of attention research. *The Korean Journal of Educational Psychology, 33*(4), 559-578.

Ji, H. (2019). *The effect of exergame on attentional and executive function of ADHD children*. Doctoral dissertation, Pusan National University.

Johnson, K. A., Robertson, I. H., Barry, E., Mulligan, A., Dáibhis, A., Daly, M., ... & Bellgrove, M. A. (2008). Impaired conflict resolution and alerting in children with ADHD: Evidence from the Attention Network Task (ANT). *Journal of Child Psychology and Psychiatry, 49*(12), 1339-1347.

Jung, I. S. (2015). *Effects of the mathematical modeling learning on the word problem solving ability and modelling process*. Doctoral dissertation, Korea National University of Education.

Kahneman, D. (1973). *Attention and effort*. New Jersey: Prentice-Hall.

Kercood, S., Zentall, S. S., & Lee, D. L. (2004). Focusing attention to deep structure in math problems: Effects on elementary education students with and without attentional deficits. *Learning and Individual Differences, 14*(2), 91-105.

Kercood, S., Zentall, S. S., Vinh, M., & Tom-Wright, K. (2012). Attentional cueing in math word problems for girls at-risk for ADHD and their peers in general education settings. *Contemporary Educational Psychology, 37*(2), 106-112.

Kibby, M. Y., Schmitter-Edgecombe, M., & Long, C. J. (1998). Ecological validity of neuropsychological tests: Focus on the California Verbal Learning Test and the Wisconsin Card Sorting Test. *Archives of Clinical Neuropsychology*, *13*(6), 523-534.

Kim, C. Y., & Kim, S. Y. (1999). Neuropsychology of attention. *Sleep Medicine and Psychophysiology, 6*(1), 26-31.

Kim, H. J., & Youn, J. J. (2013). The differences between attentional networks according to sex and reflection-impulsivity in the cognitive style of young children. *Korean Journal of Community Living Science, 24*(3), 301-312.

Kim, K. H. (2019). *The characteristic of attention networks in sluggish cognitive tempo: The effect of eye-feedback training on orienting attention in individuals with SCT.* Doctoral dissertation, Chung-Ang University.

Kim, K. W., & Kim, J. S. (2008). Development of Attention Control Scale for upper grade level elementary school students. *Korea Journal of Counseling, 51*(9), 1101-1115.

Kim, S. J., & Kim, H. N. (2006). An analysis on the influence by types and factors of word problems in the problem solving. *The Journal of Curriculum & Evaluation, 9*(2), 281-309.

Kim, S. Y., So, Y. K., Noh, J. S., Choi, N. K., Kim, S. J., & Koh, Y. J. (2003). Normative data on the Korean ADHD Rating Scales(K-ADHDRS) for parents and teacher. *Journal of Korean Neuropsychiatry Association, 42*(3), 352-359.

Kingsdorf, S., & Krawec, J. (2014). Error analysis of mathematical word problem solving across students with and without learning disabilities. *Learning Disabilities Research & Practice, 29*(2), 66-74.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155-163.

Korean Educational Psychology Association (2014). *Terminology of Educational Psychology*. Seoul: Hakjisa.

Knudsen, E. I. (2007). Fundamental components of attention. *Annual Review of Neuroscience, 30*, 57-78.

Larsen, H., van der Zwaluw, C. S., Overbeek, G., Granic, I., Franke, B., & Engels, R. C.M.E. (2010). A variable-number-of-tandem-repeats polymorphism in the dopamine D4 receptor gene affects social adaptation of alcohol use: Investigation of a gene-environment interaction. *Psychological Science, 21*(8), 1064-1068.

Lee, J. C., Kim, J. E., Kim, M. Y., Yang, J. S., Han, M. H., Kwon, H. C., ... & Kim, S. J. (2017). Test-retest reliability of attention network test scores in schizophrenia. *Korean Journal of Psychosomatic Medicine, 25*(2), 210-217.

Lee, S. H., & Park, Y. J. (1998). The neuroanatomy and psychophysiology of attention. *Sleep Medicine and Psychophysiology, 5*(2), 119-133.

Lee, Y. G. (2005). *The effects of meta-cognitive thinking on solving the word problems*. Doctoral dissertation, Dankook University.

Lewis, A. B., & Mayer, R. E. (1987). Students' miscomprehension of relational statements in arithmetic word problems. *Journal of Educational*

Psychology, 79(4), 363-371.

Long, N. M., Kuhl, B. A., & Chun, M. M. (2018). Memory and attention. *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience, 1*, 1-37.

Luszczynska, A., Diehl, M., Gutiérrez-Dona, B., Kuusinen, P., & Schwarzer, R. (2004). Measuring one component of dispositional self-regulation: Attention control in goal pursuit. *Personality and Individual Differences, 37*(3), 555-566.

MacLeod, J. W., Lawrence, M. A., McConnell, M. M., Eskes, G. A., Klein, R. M., & Shore, D. I. (2010). Appraising the ANT: Psychometric and theoretical considerations of the Attention Network Test. *Neuropsychology, 24*(5), 637-651.

Mahone, E. M., & Schneider, H. E. (2012). Assessment of attention in preschoolers. *Neuropsychology Review, 22*, 361-383.

Martin, A., Razza, R. A., & Brooks-Gunn, J. (2012). Sustained attention at age 5 predicts attention-related problems at age 9. *International Journal of Behavioral Development, 36*(6), 413-419.

Mathieu, R., Epinat-Duclos, J., Sigovan, M., Breton, A., Cheylus, A., Fayol, M., ... & Prado, J. (2018). What's behind a "+" sign? Perceiving an arithmetic operator recruits brain circuits for spatial orienting. *Cerebral Cortex, 28*(5), 1673-1684.

Mayer, R. E. (1980). Cognitive psychology and mathematical problem solving. Technical report series in learning and cognition, report no. 80-1 Retrieved from https://search.proquest.com/docview/63620638?accountid=10066

McConnell, M. M., & Shore, D. I. (2011). Mixing measures: Testing an assumption of the Attention Network Test. *Attention, Perception, & Psychophysics, 73*(4), 1096-1107.

McCowan, R. J., & McCowan, S. C. (1999). *Item analysis for criterion-referenced tests*. New York: Center for Development of Human

Services.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30-46.

Mirsky, A. F., Pascualvaca, D. M., Duncan, C. C., & French, L. M. (1999). A model of attention and its relation to ADHD. *Mental Retardation and Developmental Disabilities Research Reviews, 5*(3), 169-176.

Moosbrugger, H., & Goldhammer, F. (2014). 3.7. FAKT-II; Frankfurter Adaptiver Konzentrationsleistungs-Test II. In U. Kipman & A. Fritz (Eds.) *Psychologische Diagnostik von Aufmerksamkeits-und Konzentrationsfähigkeit im Kindergarten-und Schulalter* (pp. 68-77). Hrsg.: Österreichisches Zentrum für Begabtenförderung und Begabungsforschung.

Mukaka, M. M. (2012). A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal, 24*(3), 69-71.

Mushquash, C. J., & Bova, D. L. (2007). Cross-cultural assessment and measurement issues. *Journal on Developmental Disabilities, 13*(1), 53-65.

Musso, M. F., Boekaerts, M., Segers, M., & Cascallar, E. C. (2019). Individual differences in basic cognitive processes and self-regulated learning: Their interaction effects on math performance. *Learning and Individual Differences, 71*, 58-70.

Muth, K. D. (1992). Extraneous information and extra steps in arithmetic word problems. *Contemporary Educational Psychology, 17*(3), 278-285.

Neumann, O. (1987). Beyond capacity: A functional view of attention. *Perspectives on Perception and Action, 14*, 361-394.

Nichols, S. L., & Waschbusch, D. A. (2004). A review of the validity of laboratory cognitive tasks used to assess symptoms of ADHD. *Child Psychiatry and Human Development, 34*(4), 297-315.

Norman, D. A., & Shallice, T. (1980). Attention to action: Willed and automatic control of behavior. *Technical Report* No. 8006.

Ochsner, K. N., & Gross, J. J. (2005). The cognitive control of emotion. *Trends in Cognitive Sciences, 9*(5), 242-249.

Ogundele, M. O., Ayyash, H. F., & Banerjee, S. (2011). Role of computerized continuous performance task tests in ADHD. *Progress in Neurology and Psychiatry, 15*(3), 8-13.

Oh, H. S. (2018). *The Frankfurt Adaptive Concentration Test-II*. Seoul: Inpsyt.

Oh, K. J., & Lee, H. R. (1989). Assessment of ADHD with abbreviated Conners Rating Scale. *Korean Journal of Clinical Psychology, 8*(1), 135-142.

Passolunghi, M. C., & Lanfranchi, S. (2012). Domain specific and domain general precursors of mathematical achievement: A longitudinal study from kindergarten to first grade. *British Journal of Educational Psychology, 82*(1), 42-63.

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience, 35*, 73-89.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*(1), 25-42.

Posner, M. I., & Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. *Annual Review of Psychology, 58*, 1-23.

Pribram, K. H., & McGuinness, D. (1975). Arousal, activation, and effort in the control of attention. *Psychological Review, 82*(2), 116.

Re, A. M., Lovero, F., Cornoldi, C., & Passolunghi, M. C. (2016). Difficulties of children with ADHD symptoms in solving mathematical problems when information must be updated. *Research in Developmental Disabilities, 59*, 186-193.

Ready, R. E., Stierman, L., & Paulsen, J. S. (2001). Ecological validity of neuropsychological and personality measures of executive functions. *The*

Clinical Neuropsychologist, 15(3), 314-323.

Redick, T. S., & Engle, R. W. (2006). Working memory capacity and attention network test performance. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 20*(5), 713-721.

Robertson, L. C., Lamb, M. R., & Kight, R. T. (1988). Effects of lesions of temporal-parietal junction on perceptual and attentional processing in humans. *The Journal of Neuroscience, 8*(10), 3757-3769.

Rogers, M., Hwang, H., Toplak, M., Weiss, M., & Tannock, R. (2011). Inattention, working memory, and academic achievement in adolescents referred for attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology, 17*(5), 444-458.

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia, 42*(8), 1029-1040.

Sheridan, T. B. (2007). Attention and its allocation: Fragments of a model. In A. F. Krammer, D. A. Wiegmann, & K. Alex (Eds.), *Attention: From theory to practice* (pp. 16-26.). New York: Oxford University Press.

Sim, E. Y. (2006). *The effects of strategy training based on multiple representations on the mathematical word-problem solving*. Doctoral dissertation, Kookmin University.

Simonds, J., Kieras, J. E., Rueda, M. R., & Rothbart, M. K. (2007). Effortful control, executive attention, and emotional regulation in 7–10-year-old children. *Cognitive Development, 22*(4), 474-488.

So, Y. K., Noh, J. S., Kim, Y. S., Koh, S. G., & Koh, Y. J. (2002). The reliability and validity of Korean Parent and Teacher ADHD Rating Scale. *Journal of Korean Neuropsychiatry Association, 41*(2), 283-289.

Song, H. J. (2007). The unstability of attentional network in the schizophrenic patients. *Korean Journal of Clinical Psychology, 26*(3), 693-702.

Song, J. H. (1992). A study of the interaction between learning styles and text structures in recalling mathematical story problems. *The Korean Journal of Educational Psychology, 6*(1), 173-222.

Styles, E. A. (2006). *The psychology of attention*. New York: Psychology Press.

Valencia, R. R., & Suzuki, L. A. (2000). *Intelligence testing and minority students: Foundations, performance factors, and assessment issues* (Vol. 3). California: Sage Publications.

Van de Vijver, F., & Tanzer, N. K. (1997). Bias and equivalence in cross-cultural assessment. *European Review of Applied Psychology, 47*(4), 263-279.

Vorstenbosch, M. A., Klaassen, T. P., Kooloos, J. G., Bolhuis, S. M., & Laan, R. F. (2013). Do images influence assessment in anatomy? Exploring the effect of images on item difficulty and item discrimination. *Anatomical Sciences Education, 6*(1), 29-41.

Wallisch, A., Little, L. M., Dean, E., & Dunn, W. (2018). Executive function measures for children: A scoping review of ecological validity. *Occupation, Participation and Health, 38*(1), 6-14.

Wang, A. Y., Fuchs, L. S., & Fuchs, D. (2016). Cognitive and linguistic predictors of mathematical word problems with and without irrelevant information. *Learning and Individual Differences, 52*, 79-87.

Wang, Y. F., Jing, X. J., Liu, F., Li, M. L., Long, Z. L., Yan, J. H., & Chen, H. F. (2015). Reliable attention network scores and mutually inhibited inter-network relationships revealed by mixed design and non-orthogonal method. *Scientific Reports, 5*, 1-10.

Welford, A. T. (1952). The psychological refractory period and the timing of high-speed performance-a review and a theory. *British Journal of Psychology, 43*(1), 2-19.

Westlye, L. T., Grydeland, H., Walhovd, K. B., & Fjell, A. M. (2011).

Associations between regional cortical thickness and attentional networks as measured by the attention network test. *Cerebral Cortex, 21*(2), 345-356.

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 73*(6), 913-934.

Wu, W. (2014). *Attention*. New York: Routledge.

Yoon, C. Y., & Lee, Y. S. (2004). A normative study of Korean-Attention Deficit Hyperactivity Disorder Diagnostic Scale and developing web-based computer-assessment system. *The Korea Journal of Counselling, 5*(4), 927-939.

Zentall, S. S., Smith, Y. N., Lee, Y. B. B., & Wieczorek, C. (1994). Mathematical outcomes of attention-deficit hyperactivity disorder. *Journal of Learning Disabilities, 27*(8), 510-519.

<국문초록>

아동 주의력 네트워크 검사의 국내 생태학적 타당화

정 경 미

제주대학교 대학원 교육학전공

지도교수 송 재 홍

초등학생들의 주의력 문제를 파악하고 이를 해결하기 위한 중재 프로그램을 개발하기 위해서는 타당하고 신뢰로운 주의력 측정도구가 필수다. 이에 다양한 주의력 측정도구가 개발되어 활용되었으나 이들 도구의 심리측정학적 특징이 보다 체계적으로 평가되어야 한다는 지적이 있어 왔다. 최근 들어 신경심리학적 접근을 바탕으로 주의력이 활발히 연구되고 있다. 특히 주의력 네트워크 이론은 연구방법론적 발달과 축적된 경험적 연구 성과를 바탕으로 확립되었으며 이 이론에 기반해 주의력 네트워크 검사(Attention Network Test)와 아동 주의력 네트워크 검사(Child Attention Network Test, 이하 CANT)가 개발되었다. 국외에서 개발된 검사를 국내 샘플을 대상으로 적용했을 때 검사 결과에 문화적 차이로 인한 편향이 개입될 수 있다. 따라서 국외에서 개발된 측정 도구를 사용하기 위해서는 우선 국내 샘플을 대상으로 도구의 심리측정학적 특징을 검증해야 한다. 또한 신경심리학적 측정도구는 통제된 실험실 환경에서 제한된 자극에 대한 반응의 특정 측면만 표집해 피검사의 인지적 특징을 추론하므로 검사 결과가 일상 생활에서의 문제나 수행에 일반화될 수 있는지 확인하는 과정도 필요하다. 이에 이 연구에서는 CANT의 심리측정학적 특징이 원래 개발되어진 문화권이 아닌 국내에서 양호한지와 통제된 실험실적 환경에서 얻어진 CANT 검사 결과가 학습 장면 특히 수문장제 수행을 예측하는지를 분석해 CANT의 생태학적 타당도를 검증하고자 한다. 이를 통해 초등학생 주의력 연구

의 기반을 마련하고 주의력이 학업성취에 작동하는 방식에 대한 시사점을 찾아 초등학생의 성공적인 학교생활에 기여하는 것이 이 연구의 목적이다. 연구 목적을 달성하기 위해 연구 Ⅰ에서는 아동 주의력 네트워크 검사의 문항양호도와 검사 신뢰도 및 타당도를 국내 샘플을 대상으로 분석했다. 연구 Ⅱ에서는 아동 주의력 네트워크 검사의 생태학적 타당도를 임상장면이 아닌 학교의 수문장제 문제 해결 능력과의 상관을 중심으로 분석했다.

연구 Ⅰ의 연구문제를 해결하기 위해 OO 지역 초등학교 1-6학년 학생 71명을 표집해 아동 주의력 네트워크 검사를 실시했다. 이 학생 중 4-6학년의 학생 35명은 10-14일 후 다시 CANT를 수검했으며 공인타당도 검증을 위해 FACT-Ⅱ-SR 검사도 받았다. 또한 이 학생들의 학부모와 담임교사도 연구에 참여해 대상 학생들의 가정과 학교에서의 부주의 관련 행동을 KADHDDS와 K-ADHDRS-T를 활용해 평정했다. 아동 주의력 네트워크 검사의 문항양호도는 문항변별도와 문항타당도 측면에서 분석했고 검사 신뢰도는 Cronbach's α, 반분신뢰도, 검사-재검사 신뢰도 측면에서 분석했다. 검사 타당도는 공인타당도 측면에서 분석했다. 문항양호도 검증 결과 아동 주의력 네트워크 검사의 문항은 정확도와 반응 시간 측면에서 적합한 변별력을 갖추고 있음을 확인했다. 또한 신호 조건과 플랭커 조건이 문항별 정확도와 반응시간에 유의한 효과를 보이고 있는 것으로 나타나 문항이 타당한 것으로 판단되었다. CANT의 신뢰도는 측정치에 따라 다른 분포를 보였다. 정확도의 Cronbach's α는 .90, 반분신뢰도는 .52-.65, 검사-재검사 신뢰도는 .17로 나타났다. 반응시간의 Cronbach's α는 .99, 반분신뢰도는 .93-.94, 검사-재검사 신뢰도는 .90이었다. 각성 네트워크, 정향 네트워크, 관리통제 네트워크의 반분신뢰도는 .05-.39의 분포를 보였으며 검사-재검사 신뢰도는 .16-.46의 분포를 보였다. 공인타당도 검증 결과 CANT의 정향 네트워크, 관리통제 네트워크, 정확도, 반응시간이 수행 기반 검사인 FACT-Ⅱ-SR 구인과 유의미한 상관관계에 있는 것으로 확인되었으며 평정형 검사의 구인과는 정향 네트워크, 관리통제 네트워크, 정확도가 유의미한 상관을 보이고 있는 것으로 나타났다. 이는 아동 주의력 네트워크 검사가 국내에서도 타당한 도구나 신뢰도 향상을 위한 추가적 노력이 필요함을 시사한다.

연구 Ⅱ 문제를 해결하기 위해 OO 지역 초등학교 5-6학년 학생 31명을 표집해 아동 주의력 네트워크 검사와 수문장제 검사를 실시했다. 이후 주의력 네트워크 검사가 수문장제 수행 능력을 예측하는지 알아보기 위해 피어슨 상관분석, 위계적 회귀분석, 수행패턴 시각화를 실시했다. 연구 결과 아동 주의력 네트워크 검사의 정확도는 수문장제의 정확도를 유의하게 설명하고 있으며 수학 계산 능력을 매개로 수문장제의 반응 시간도 유

의하게 예측하고 있음이 확인되었다. 또한 각성 네트워크, 정향 네트워크, 관리통제 네트워크가 불필요한 정보나 불일치 관계 용어를 포함한 수문장제 수행의 분산을 유의하게 설명하지는 못했으나 각 네트워크 점수는 특정 유형의 수문장제 수행 속도 혹은 정확도와 유의한 상관을 갖고 있는 것으로 나타났다. 또한 주의력 네트워크 검사에서의 수행 양상이 정확도와 속도 측면에서 수문장제 해결 수행 양상과 유사한 것으로 나타났다. 이는 아동 주의력 네트워크 검사 결과가 수문장제 해결에 일반화될 수 있으며 따라서 CANT가 생태학적으로 타당한 도구임을 의미한다.

이 연구는 아동 주의력 네트워크 검사의 생태학적 특징을 분석함으로써 추후 주의력 연구의 기반을 마련하는 데 공헌했다. 연구 결과에 따르면 아동 주의력 네트워크 검사는 타당하고 변별력을 갖춘 문항으로 구성되었으며 다른 주의력 검사와 상관을 가진 타당한 검사도구다. CANT의 정확도와 반응시간은 수용할 만한 신뢰도를 보이고 있으나 세 가지 네트워크 점수의 신뢰도를 향상시키기 위한 노력이 필요한 것으로 나타났다. 아울러 아동 주의력 네트워크 검사 결과는 수문장제 수행 능력을 유의하게 예측하는 것으로 나타났다. 즉 아동 주의력 네트워크 검사는 국외 뿐 아니라 국내에서, 임상장면 뿐 아니라 학교장면에서도 타당한 주의력 측정도구이다. 후속 연구에서는 아동 주의력 네트워크 검사를 지속적으로 타당화해야 하며 다양한 방법을 통해 신뢰도를 향상시켜야겠다. 나아가 아동 주의력 네트워크 점수와 수문장제 문제 해결능력 간의 관계를 바탕으로 초등학생들의 성공적인 학교생활에 기여하는 주의력 변인에 대한 연구가 다양한 학문 영역에서 후속되어야 하겠다.

Appendix

Appendix A.

Korean Attention Deficit Hyperactivity Disorder Diagnostic Scale (KADHDDS, Inattention Subscale)

Child's Name		Grade	
Completed by		Date	

This questionnaire is to check the attention-related behaviors of your child. Rate your child's behavior, as described in the box below.

0 : Absence of symptoms (rarely or does not show the behavior)
1 : Mild symptoms (Sometimes but not severely shows the behavior)
2 : Severe symptoms (Frequently and severely shows the behavior)

Items	Absence of symptoms	Mild symptoms	Severe symptoms
1. Does not seem to listen to others.	0	1	2
2. Is often forgetful in daily activities (appointment, homework.)	0	1	2
3. Does not follow through on instructions and requests.	0	1	2
4. Has difficulty planning.	0	1	2
5. Has difficulty in completing a planned task.	0	1	2
6. Often makes careless mistakes.	0	1	2
7. Has difficulty paying attention to plays or activities and cleaning up.	0	1	2
8. Pays attention only for a short period.	0	1	2
9. Is easily distracted by extraneous stimuli	0	1	2
10. Is easily irritated by and gives up challenging tasks.	0	1	2
11. Has difficulty in concentrating on a task.	0	1	2
12. Has difficulty completing a task.	0	1	2
13. Often loses things.	0	1	2

Appendix B.

Korean Teacher ADHD Rating Scale
(K-ADHDRS-T, Inattention Subscale)

Student Name		Grade	
Completed by		Date	

This questionnaire is to check the attention-related behaviors of your student. Circle the number that best describes this students' school behavior over the past 6 months (or since the beginning of the school year).

Items	Never or rarely	Sometimes	Often	Very often
1. Fails to give close attention to details or makes careless mistakes in schoolwork.	0	1	2	3
2. Has difficulty sustaining attention in tasks or play activities.	0	1	2	3
3. Does not seem to listen when spoken to directly.	0	1	2	3
4. Does not follow through on instructions and fails to finish work.	0	1	2	3
5. Has difficulty organizing tasks and activities.	0	1	2	3
6. Avoids tasks (e.g., schoolwork, homework) that require sustained mental effort.	0	1	2	3
7. Loses things necessary for tasks or activities.	0	1	2	3
8. Is easily distracted.	0	1	2	3
9. Is forgetful in daily activities.	0	1	2	3

Appendix C.

Mathematical Computation Test

<Practice> 7 + 2 =

1. 8 ÷ 2 =

2. 83 + 12 =

3. 30 x 3 =

4. 37 - 14 =

5. 150 ÷ 5 =

6. 8 + 25 =

7. 3500 - 300 =

8. 25 x 7 =

9. 84 ÷ 7 =

10. 7 x 7 =

Appendix D.

Mathematical Word Problem Test

<Practice> Jisu ate six apples last month. Segyeong ate 2 apples less than Jisu. How many apples did Segyeong eat last month?

1. Dongho runs 6km per week for practice. He runs three times longer than Sooyeong runs per week. How long does Sooyeong run per week?

2. There are 98 boys in grade 5. The boys' number is 4 less than the number of girls. How many girls are there in grade 5?

3. There are 20 croakers in a pack. There are four packs of croakers. How many croakers are there in total?

4. Yesterday Jaeheon picked 84 pears in the orchard and put them in 7 boxes. This number of pears is 19 less than the number of pears he picked today. How many pears did Jaeheon pick today?

5. Soojeong has 26 marbles. Yeongseon has 8 more marbles than Soojeong, and Hyeonhee has 15 marbles less than Soojeong. How many marbles does Hyeonhee have?

6. The circumference of the notebook is 36cm. If the sum of the width is 14cm, what is the sum of the length?

7. The convenience store sells 120 bottles of water per day. This is as twice

as the number of bottles of water that the other supermarket sells per day. How many bottles of water does the supermarket sell per day?

8. The house is 7km away from the library, and the library is 14km away from the school. How far is the house from the school passing the library?

9. A bakery shop sells a big box which can contain 25 cookies at 2,400 won. The big box is 400 won more expensive than the small box which can contain 20 cookies. How much is a small box?

10. The older brother's wooden stick is 94cm long. This stick is 16cm longer than the younger brother's wooden stick. How long is the younger brothers' stick?

11. There are 8 small candies and 24 dry cookies in a snack pack. How many dry cookies are there in 6 snack packs?

12. A bottle of strawberry juice weighs 6kg. This is twice as heavy as a bottle of orange juice. There is 20g of orange in a bottle of orange juice. How heavy is a bottle of orange juice?

13. Four children equally shared the 72m long wire. How long is the wire that each child got?

14. Jinsoo weighs 21kg, and his father weighs 40kg more than Jinsoo. Jinsoo's brother weighs 32kg more than Jinsoo. How much does Jinsoo's father weigh?

15. The teacher gave 32 students 4 oranges and 2 apples for each person. If there is no orange and apple left, how many oranges did the teacher have at first?

16. There are a 25cm^2-area equilateral triangle and a 16cm^2-area square. How long is a straight side of the square?

17. Jaehyeop picked 84 pears in the orchard yesterday. The number of pears he picked today is 19 more than the number of pears he picked yesterday. How many pears did he pick today?

18. Eunjoon gave 8 friends 2 pencils for each, and then there was no pencil left. How many pencils did Eunjoon have in the beginning?

19. Miyeong and Yeongho participated in the marathon race to donate money for people in need. Miyeong donated 2,500 won per 1km-long running, and Yeonghoo donated 3,000 won per 1 km-long running. Eventually, Yeongho donated 15,000 won after the marathon. How long did he run in the marathon?

20. “A” oil station sells 1L of oil at the price of 1,547 won. This price is 5 won cheaper than the price at “B” oil station. How much is 1L of oil at B oil station?