

The Construction and Validation of the Scientific Thinking Disposition Inventory

Jeng-Fung Hung, Kun Lee, Men-An Pan
National Kaohsiung Normal University, Kaohsiung, Taiwan

Fu-Pei Hsieh
Kuang-Hua Primary School, Kaohsiung, Taiwan

This study aimed to develop a scientific thinking disposition inventory (STDI) with good reliability and validity. The construction was a two-stage process. In the first stage, the draft of the STDI with 60 items was completed based on the related literature. Next, a pilot study with 309 students as the subjects was conducted for removing the items with unsatisfying item analysis results from the STDI draft. The modified STDI contained five subscales and 44 items. Then, reliability analyses were performed with the subscales. The obtained Cronbach's α -value were between 0.653 and 0.830, of which the STDI was 0.952, meaning that the reliability of this inventory was good. The result of the factor analysis showed that only one factor—scientific thinking disposition, could be extracted, meaning that when applying the STDI to research. The results must be explained based on the total scores and it is not appropriate to discuss any of the subscales separately. The second stage aimed to explore the differences in scientific creativity and scientific reasoning among students from different scientific thinking disposition groups with 171 students as the research subjects. The purpose was to examine the construct validity of the inventory. It was found that the differences in scientific reasoning test scores and scientific creativity test scores between the groups of high STDI scores and that of low STDI scores were significant ($p < 0.001$), showing that the construct validity of this inventory was high.

Keywords: habits of mind, scientific thinking disposition, scientific thinking disposition inventory (STDI), thinking disposition, thinking patterns

Introduction

Individual thinking patterns may be influenced by value systems in science and these thinking patterns are shared by science traditions. Thus, they have become scientific thinking dispositions (Settlage & Southerland, 2007). These thinking dispositions are not spontaneous. They are passed down from one generation of scientists to the next. Students who just started to learn science must learn behaviors conforming to these thinking dispositions in the process of becoming members of science culture. The meanings of these thinking

Jeng-Fung Hung, Ph.D., professor, Graduate Institute of Science Education and Environmental Education, National Kaohsiung Normal University.

Men-An Pan, Ph.D. student, Graduate Institute of Science Education and Environmental Education, National Kaohsiung Normal University.

Kun Lee, Ph.D. student, Graduate Institute of Science Education and Environmental Education, National Kaohsiung Normal University.

Fu-Pei Hsieh, Ph.D., teacher, Kuang-Hua Primary School.

dispositions can be explored from the literature on scientists.

According to the studies by McGrayne (1998) and Louis, Holdsworth, Anderson, and Campbell (2007), successful scientists can achieve great things because of not only their high intelligence quotient (IQ), but also their high curiosity, strong motivation, persistence, and hard working. The characteristics of outstanding scientists include: (a) having love for and an interest in scientific research; (b) being passionate and confident; (c) having the scientific attitude of always getting to the bottom of an affair; (d) spending a long time and efforts thinking; (e) being persistent; (f) working hard; (g) not being afraid of frustration; (h) learning from the wise; (i) doing research with others; (j) being innovative; (k) always learning new knowledge; (l) pursuing high precision; (m) being sensitive; and (n) being capable of organizational thinking.

The American Association for the Advancement of Science (AAAS, 1989) published a book *Science for All Americans* with their Project 2061. In this book, there is an independent chapter titled "Thinking Disposition," stressing that students should develop some thinking dispositions in relation to the values, attitudes, and skills in science culture during their school years. According to this chapter, the most important scientific thinking dispositions include curiosity, openness to new ideas, and skepticism. The standard B from *National Science Education Standards* (National Research Council [NRC], 1996) also clearly specified that students should be cultivated to possess the above characteristics in science culture.

Literature Review

In order to give a more complete definition of scientific thinking dispositions, this study also referred to the literature of thinking dispositions and habits of mind. Siegel (1999) suggested that a thinking disposition is a person's true tendency or habit of thinking in a certain way in a certain context. It is thinkers' asset and helps thinkers to do quality thinking. It is beneficial for quality thinking. Stanovich and West (1997) believed that a thinking disposition is a very robust mental mechanism for producing certain behavioral tendencies and strategies, such as carefully assessing new evidences against the dominant idea, spending a lot of time solving a problem before giving up, and carefully considering others' perspectives before forming a viewpoint. Perkins and Tishman (1998) indicated that a thinking disposition is about showing a consistent and persistent thinking behavior in different thinking situations in the long term and is composed of three elements: capability, sensitivity, and tendency. The capability is the potential and skill to conduct a certain behavior, sensitivity is alertness regarding the behavior suitable for a certain circumstance, and tendency is the willingness to actually do something in a certain aspect. Stanovich (2001) stressed that thinking dispositions help to clarify one's own goals and the value of knowledge and to show the tendency toward pragmatism and self-discipline. This definition reflects the psychological structure at the rational level. What kinds of thinking dispositions are quality ones? The seven dispositions proposed by Tishman, Jay, and Perkins (1993) were the most representative quality thinking dispositions as the following:

1. The disposition to be broad and adventurous;
2. The disposition to wonder, find problem, and investigate;
3. The disposition to build explanations and understandings;
4. The disposition to make plans and be strategic;
5. The disposition to be intellectually careful;
6. The disposition to seek and evaluate reasons;
7. The disposition to be metacognitive.

The discussions on habits of mind are associated with the meanings of a thinking disposition. Costa and Kallick (2000) defined habits of mind as a pattern of intellectual behaviors which can lead to actions with achievements and a characteristic of problem solvers with a quick wit. Perkins et al. (2000) and Costa and Kallick (2000) believed that even though habits of mind and thinking dispositions are different in amounts and names, they are similar in the psychological aspect, as they both focus on open-mind, rationality, curiosity, and metacognitive reflective-thinking. The 16 habits of mind are listed as below:

1. Persisting;
2. Thinking and communicating with clarity and precision;
3. Managing impulsivity;
4. Gathering data through all senses;
5. Listening with understanding and empathy;
6. Creating, imagining, and innovating;
7. Thinking flexibly;
8. Responding with wonderment and awe;
9. Thinking about thinking (metacognition);
10. Taking responsible risks;
11. Striving for accuracy;
12. Finding humor;
13. Questioning and posing problems;
14. Thinking interdependently;
15. Applying past knowledge to new situations;
16. Remaining open to continuous learning.

It was found through the literature review above that scientific thinking dispositions are almost the same as habits of mind if the domain-specific characteristics are not taken into consideration. Thus, based on the discussions above, this study decided to develop the scientific thinking disposition inventory (STDI) based on 10 thinking dispositions, namely, persistence, managing impulsivity, listening with understanding and empathy, thinking flexibly, reflective thinking, striving for accuracy, questioning and posing problems, taking a critical perspective, thinking interdependently, and organizational thinking. And science terms, such as scientific problem and experiment report, were used in item descriptions.

This study examined the construct validity using the relationship between scientific creativity and scientific reasoning. According to the viewpoint by Fraenkel and Wallen (2006), there were three steps to obtain a proof for construct validity:

1. Clearly defining the variable being measured;
2. Creating a hypothesis based on the theory of that variable, usually in this form—Those who score high on this variable would show a certain behavior under certain circumstance, while those who score low would show another behavior under the same circumstance;
3. Testing the hypothesis by using logical inference and empirical methods.

From this perspective, it was possible to explore the significance of the differences in behaviors under the influence of certain scientific thinking dispositions between the two groups of students with high and low STDI scores, for examining the construct validity of the STDI.

Based on the above, this study decided to examine the construct validity of the STDI using two variables—scientific creativity and scientific reasoning. Below is a detailed explanation.

Hu and Adey (2002) believed that, through processes including imagination and thinking, scientific creativity is to produce technical products with fluency, flexibility, and originality or products of science knowledge, science phenomena, or science problems. And fluency, flexibility, and originality are related to the science thinking disposition “thinking flexibly.” In the aspect of science reasoning, according to the definition by Lawson (1995), there were seven types of science reasoning capability include: classification, conservation, the identification and control of variables, proportional, combinatorial, probabilistic, and correlational reasoning. These types of reasoning capability are related to the science thinking dispositions “striving for accuracy,” “organizational thinking,” and “reflective thinking.” A science thinking disposition is the persistence of continuous thinking behavior and tendency in different scientific thinking situations in the long term (Perkins & Tishman, 1998). Developing good thinking dispositions helps to achieve more mature and wiser thinking. (Costa & Kallick, 2000)

According to the above statement, it can be reasonably inferred that good science thinking dispositions lead to improvements in scientific creativity and scientific reasoning performances.

In sum, this study’s hypothesis is that the scientific creativity and scientific reasoning performances of those with good scientific thinking dispositions are better than the performances of those with bad scientific thinking dispositions. Thus, this study used the STDI, the scientific creativity test, and the scientific reasoning test to verify the hypothesis that those with high STDI scores would score high in the scientific creativity test and the scientific reasoning test, while those with low STDI scores would score low in the scientific creativity test and the scientific reasoning test, to examine the construct validity of the STDI.

Research Design

This study aimed to develop a STDI with good reliability and validity. It also aims to explore whether the modified inventory could be used to effectively examine the differences in scientific creativity and scientific test scores between students of different scientific thinking dispositions. The development process of the STDI and the research samples are discussed below.

The Development Process of the STDI

There were two main stages of the development process of the STDI. In the first stage, the 10 thinking dispositions were used as constructs. The theoretical explanations by Tishman, Jay, and Perkins (1993) and Costa and Kallick (2000) were used as references for the descriptions of the items of the STDI. The draft contained 60 items corresponding to the scientific thinking dispositions. Then, two scholars from the field of science education and three junior high school natural science teachers were invited to offer their opinions to modify the draft to improve the content validity. Before the pre-test, three eighth graders who were not among the research subjects were asked to read the inventory. The inventory was modified based on their opinions to make sure that students could understand it and to improve the face validity.

After the STDI draft was completed, item analyses were conducted to explore the reliabilities of the items individually. The methods used included comparisons of extreme groups and homogeneity tests. The former could be used as a discrimination index. With the significance level being achieved, the higher the critical value (t) was, the higher the corresponding item discrimination. The later showed whether the psychological traits being measured by the items were similar. The correlation between the items and the complete inventory should be above 0.30 (Murphy & Davidshofer, 2001; McIntire & Miller, 2007).

Inappropriate items were removed based on the item analysis results of the draft. Then, the reliability of

the whole inventory was assessed through reliability analyses. If the reliability value of a subscale was lower than 0.60, the descriptions of its items were modified or new items were added. If this criterion was met, the correlation coefficients between the subscales were calculated. Coefficients lower than 0.2 were considered too low and those higher than 0.8 were considered too high. In these cases, further discussions were required to determine whether the corresponding subscales should be removed or combined. Then, exploratory and confirmatory factor analyses were performed to examine the corresponding validity (Arbuckle & Wothke, 1999; Murphy & Davidshofer, 2001; McIntire & Miller, 2007).

In the second stage, according to the viewpoint proposed by Fraenkel and Wallen (2006), the three steps to obtain proofs for construct validity were conducted. After the STDI was finalized, different research samples were used to conduct the STDI survey, the scientific reasoning test scores, and the scientific creativity test. It also used to verify the hypothesis that those with high STDI scores would score high in the scientific creativity test and the scientific reasoning test, while those with low STDI scores would score low in the scientific creativity test and the scientific reasoning test.

Samples

According to the findings of the study by Perkins et al. (2000), the re-test reliability values for perception and judgment of ninth graders were 0.81 and 0.72, respectively. Those of fifth graders were 0.61 and 0.68, respectively, showing that younger students were more likely to have different performances in thinking dispositions. Because this study aimed to develop a STDI junior high school students with rather robust performances in thinking dispositions were considered as candidates for the research subjects. The researcher of this study consulted some peer teachers and learned that ninth graders were not able to work with this study as they had to prepare for the senior high school entrance exam. Thus, the eighth graders were selected as the research subjects.

Stage I. The schools for research sampling included the large schools in the metropolitan area and the medium-sized schools in the small metropolitan area in Kaohsiung. These schools were all normal class-grouping schools. A total of 309 eighth graders from 11 classes participated in the pre-test.

Stage II. The schools for research sampling were the same as those in Stage I. One hundred and seventy-one eighth graders who were not among the participants in Stage I were selected as the research subjects.

Instruments

This study adopted the scientific creativity test developed by Hu and Adey (2002) and modified by Chen (2004), to measure the students' scientific creativity shown while learning science. The items were categorized into four groups: science knowledge, science question, technology product, and scientific phenomenon. The reliability of the modified test was 0.630, while its correlation coefficient with the originality measure in the Torrance tests of creative thinking was 0.452. The reliability and criterion-related validity were both good and the content was readable for Taiwanese students (Chen, 2004). The scientific reasoning test used in this study is the test modified by Lawson in 1995 with 12 questions groups from seven categories: conservation of mass, conservation of volume, ratio, variable control, probability, combination, and relationship inference. The Cronbach α -value of the Chinese version of this test was 0.83. Although some scientific concepts, such as capillarity, gravity, and ratio, were involved in the scientific creativity test and the scientific reasoning test, the eighth graders had already learned these concepts from the previous science courses.

Findings and Discussions

The research findings are summarized below based on the implementation stages of this study.

Stage I

The content of the draft of the STDI. The literature review above shows that scientific thinking dispositions include the following 10 habits: persistence, managing impulsivity, listening with understanding and empathy, thinking flexibly, reflective thinking, striving for accuracy, questioning and posing problems, taking a critical perspective, thinking interdependently, and organizational thinking. Thus, the researcher used the definitions of the behaviors as the constructs and designed six items for each of the constructs. For the purpose of applying the inventory to large samples and convenience of scoring, 5-point Likert scale was used for the tool with five options from “Very likely” to “Very unlikely.” The meanings and items of the subscales are summarized below (see Table 1).

Table 1

The Items Used in the Draft of the STDI

Subscale	No.	Description
Persistence	1	When learning science, no matter how difficult the problem I encounter is, I would persist in achieving my goal and I would never give up easily.
	11	When encountering a difficult scientific question that I cannot solve, I often just quit and give up. (r)
	21	When exploring a scientific question, even in a chaotic and confusing situation, I would not give up and would hold on till the last moment.
	31	When exploring a scientific question, I would not back down no matter how many frustrations come along.
	41	When I encounter a scientific question that I cannot solve, I would just come up with an answer and get it over with. (r)
	51	If I cannot complete the exploration of a scientific question when I try for the first time, I would keep on trying until I succeed.
Managing impulsivity	2	When exploring a scientific question, I would not make any judgment until the right moment.
	12	When exploring a scientific question, if I can think of the answer, I would say it out loudly immediately. (r)
	22	When exploring a scientific question, I would patiently confirm the answer in my mind before giving it.
	32	When exploring a scientific question, I would spend some time assessing other alternatives and possible results.
	42	When exploring a scientific question, I often take actions right after hearing a suggestion of coming up with a strategy. (r)
	52	When exploring a scientific question, I would conduct experiments over and over until I am satisfied with the results before writing down my conclusion.
Listening with understanding and empathy	3	When discussing a scientific question with others, I would stop my own thinking for a moment to carefully consider others' perspectives.
	13	When discussing a scientific question with others, I can learn diversified perspectives from others.
	23	When discussing a scientific question with others, I often ponder on what I am going to answer in my mind. (r)
	33	When discussing a scientific question with others, I would extend their perspectives and clarify with them.
	43	When discussing a scientific question with others, I would interpret their perspectives accurately.
53	When discussing a scientific question with others, I am better at controlling my own trains of thought than at understanding others'. (r)	
Thinking flexibly	4	When exploring a scientific question, I would come up with various ideas based on data from multiple sources.
	14	When exploring a scientific question, I do not know what to do if my first solution does not work. (r)
	24	When exploring a scientific question, I would come up with different methods and solutions.
	34	When exploring a scientific question, I would think from different perspectives.
	44	When exploring a scientific question, I could accept reasoning that is completely opposite to my viewpoint.
54	When exploring a scientific question, I am not good at thinking about the question from different perspectives. (r)	

(Table 1 to be continued)

Reflective thinking	5	When exploring a scientific question, I would monitor my own thought to find my bias.
	15	When exploring a scientific question, all I care about is to complete the corresponding experiment or assignment and I rarely think about "Why I do what I do." (r)
	25	When exploring a scientific question, I would think about strategies for future improvement after the exploration is completed.
	35	When exploring a scientific question, I rarely question my own thinking strategies. (r)
	45	When exploring a scientific question, I would keep on reviewing my strategies to make sure they are appropriate.
	55	When exploring a scientific question, I rarely evaluate the effects after the exploration is completed.
Striving for accuracy	6	When exploring a scientific question, I would review the goal to be achieved.
	16	There are often mistakes in science experiment reports I submit. (r)
	26	When exploring a scientific question, I would review the laws that should be followed.
	36	When exploring a scientific question, I would review the conditions that should be complied with.
	46	If I have to choose between submitting a science experiment report on time and making sure there is no mistake in the report, I would choose the former. (r)
	56	When learning science, I would look over my assignment again after it is completed to try to correct all mistakes if possible without being reminded.
Questioning and posing problems	7	When learning science, I often think, "If ..., what would happen?"
	17	When learning science, I often think, "What is the connection between these things?"
	27	When learning science, I often think, "What is the cause of this relationship?"
	37	When exploring a scientific question, I would also look for issues that can be further looked into after making a conclusion.
	47	If there is something abnormal in the result of a science experiment report, I often choose to ignore it. (r)
	57	When learning science, I would ask questions actively in order to obtain new information.
Taking a critical perspective	8	When exploring a scientific question, I am not good at deciding whether an explanation is better than the other. (r)
	18	When exploring a scientific question, I would take the evidence against my viewpoint into consideration.
	28	When exploring a scientific question, I would compare the advantages and disadvantages between different solutions.
	38	When exploring a scientific question, I am not good at determining which information is related to the question and which is not. (r)
	48	When exploring a scientific question, I would check the consistency between inferences.
	58	When exploring a scientific question, I would make sure my conclusion is based on current evidences and reasonable reasons.
Thinking interdependently	9	During a group discussion, I would comment on another person's idea through a supportive way.
	19	During a group discussion, I would provide positive opinions and opinions about improvements that can be made.
	29	During a group discussion, I would hold on to my own view and make sure others agree with me. (r)
	39	During a group discussion, I would discuss with others about different views.
	49	During a group discussion, I am not good at expressing my ideas and views. (r)
	59	During a group discussion, other people's views often help me to come up with a new question or a new thinking direction.
Organizational thinking	10	When exploring a scientific question, I would organize various ideas in a logical way.
	20	When exploring a scientific question, I am not good at planning the steps for exploration. (r)
	30	When exploring a scientific question, I would divide the question into several parts in an organized way.
	40	I am used to explore scientific questions using the thinking method of "taking actions for every thought that comes to mind." (r)
	50	When exploring a scientific question, I would do it step by step and logically.
	60	I would explore scientific questions with plans and methods.

Note. Items with an "r" are reverse coded items.

The item analysis of the draft of the STDI. The item analyses for the STDI were performed using a statistical analysis software package—statistic package for social science (SPSS) for Windows 17.0. The results are summarized in Table 2. The scores of Items 12, 23, 29, 40, 42, and 46 of the low STDI score group were higher than those of the high STDI score group, which was obviously not reasonable. The critical values of Items 8, 35, and 53 were not significant. The homogeneity test values for Items 16, 20, 47, and 49 were too low, meaning that these items were not ideal and should be removed.

Table 2

The Item Analysis of the 60 Items of the Draft of the STDI

No. in the original inventory	Comparison of extreme groups		Homogeneity test			No. after the items to be removed
	Critical value	<i>p</i>	Correlation with the inventory	Correlation with the subscale	Reliability value after the items to be removed	
1	11.397	0.000	0.575	0.510	0.921	1
2	5.238	0.000	0.294	0.231	0.923	2
3	7.109	0.000	0.412	0.446	0.922	3
4	10.658	0.000	0.609	0.512	0.921	4
5	9.925	0.000	0.535	0.338	0.921	5
6	9.797	0.000	0.559	0.338	0.921	6
7	8.713	0.000	0.498	0.504	0.922	7
8	1.149	0.252	0.026	-0.031	0.925	removed
9	5.709	0.000	0.286	0.281	0.923	8
10	11.903	0.000	0.564	0.416	0.921	9
11	4.621	0.000	0.376	0.387	0.922	10
12	-1.734	0.085	-0.094	-0.056	0.926	removed
13	9.730	0.000	0.492	0.373	0.922	11
14	5.575	0.000	0.395	0.310	0.922	12
15	4.163	0.000	0.305	0.203	0.923	removed
16	2.107	0.037	0.134	0.000	0.924	removed
17	9.383	0.000	0.545	0.553	0.921	13
18	9.911	0.000	0.514	0.371	0.921	14
19	9.542	0.000	0.533	0.422	0.921	15
20	2.643	0.009	0.249	0.141	0.923	removed
21	12.477	0.000	0.646	0.627	0.920	16
22	10.288	0.000	0.552	0.321	0.921	17
23	-9.046	0.000	-0.518	0.419	0.930	removed
24	11.234	0.000	0.598	0.530	0.921	18
25	10.737	0.000	0.584	0.311	0.921	19
26	10.934	0.000	0.604	0.391	0.921	20
27	10.562	0.000	0.604	0.626	0.921	21
28	8.405	0.000	0.502	0.349	0.922	22
29	-2.493	0.014	-0.134	-0.142	0.926	removed
30	8.893	0.000	0.527	0.417	0.921	23
31	11.091	0.000	0.657	0.597	0.920	24
32	10.441	0.000	0.612	0.212	0.921	25
33	9.389	0.000	0.469	0.339	0.922	26
34	11.418	0.000	0.629	0.576	0.921	27
35	0.349	0.727	0.024	0.122	0.925	removed
36	9.650	0.000	0.505	0.344	0.922	removed
37	12.131	0.000	0.642	0.606	0.921	28

(Table 2 to be continued)

38	2.265	0.025	0.248	0.130	0.923	29
39	8.806	0.000	0.476	0.316	0.922	30
40	-1.454	0.148	-0.165	-0.184	0.927	removed
41	7.180	0.000	0.469	0.403	0.922	31
42	-4.233	0.000	-0.319	-0.219	0.927	removed
43	8.199	0.000	0.506	0.366	0.922	32
44	6.361	0.000	0.383	0.277	0.922	33
45	9.302	0.000	0.579	0.348	0.921	34
46	-1.077	0.283	-0.075	-0.108	0.926	removed
47	4.905	0.000	0.337	0.191	0.923	removed
48	10.852	0.000	0.635	0.457	0.921	35
49	2.926	0.004	0.222	0.084	0.924	removed
50	10.679	0.000	0.611	0.449	0.921	36
51	12.253	0.000	0.639	0.565	0.920	37
52	9.946	0.000	0.595	0.284	0.921	38
53	1.881	0.062	0.109	0.093	0.924	removed
54	2.766	0.006	0.273	0.148	0.923	39
55	2.559	0.011	0.201	0.205	0.924	removed
56	7.910	0.000	0.540	0.422	0.921	40
57	11.644	0.000	0.635	0.522	0.921	41
58	9.991	0.000	0.578	0.358	0.921	42
59	12.229	0.000	0.620	0.396	0.921	43
60	11.414	0.000	0.647	0.441	0.920	44

Note. Items marked in bold were to be removed.

Table 3

The α -Value of the STDI Subscales Before and After Merging and the Overall Inventory

Before merging		After merging	
Subscale and No.	α	Subscale (Original No.)	α
1. Persistence	0.766	Persistence and managing impulsivity (1, 2)	0.742
2. Managing impulsivity	0.270		
3. Listening with understanding and empathy	0.603	Thinking interdependently (3, 9)	0.528
4. Thinking flexibly	0.653	Thinking flexibly (4)	0.653
5. Reflective thinking	0.494	Metacognition (5, 6, 10)	0.762
6. Striving for accuracy	0.437		
7. Questioning and posing problems	0.753	Taking a critical perspective (7, 8)	0.797
8. Taking a critical perspective	0.500		
9. Thinking interdependently	0.425		
10. Organizational thinking	0.502		
Overall inventory	0.924		

However, the researcher found that the reliability of the inventory was 0.924, while those of subscales 2, 5, 6, 8, 9, and 10 were merely between 0.270 and 0.502 (see Table 3), which were obviously too low, meaning that the items of these subscales needed to be re-arranged.

First of all, according to the definitions in the literature review, subscales of similar meanings were merged. For example, "persistence" means working hard to hold on to the last moment without giving up and

“managing impulsivity” means thinking before acting, keeping the mind calm, pondering deeply over, and being careful and cautious. The definitions of these two subscales are rather similar compared to other subscales. Thus, these two subscales were merged into the “persistence and managing impulsivity” subscale.

“Listening with understanding and empathy” was defined as trying to understand others, putting in some efforts to take others’ ideas and perspectives into consideration, and carefully appreciating others’ views and goodwill, while “thinking interdependently” was defined as working together and truly cooperate with and learn from others in a mutually beneficial situation. The definitions of these two subscales are rather similar compared to other subscales. Thus, these two subscales were merged into the “thinking interdependently” subscale.

“Thinking reflectively” was defined as understanding one’s own thinking patterns, paying attention to one’s own thoughts, strategies, feelings, and conducts, and thinking about the influences of these conducts on others, while “striving for accuracy” was defined as developing the desire to pursue accuracy, precision, and skills. “Organizational thinking” was defined as applying previously learned knowledge to other scenarios. The definitions of these three subscales are rather similar compared to other subscales. Thus, these three subscales were merged into the “metacognition” subscale.

“Questioning and posing problems” was defined as developing the attitude of seeking knowledge and asking questions, to look for not only problems but also answers to these problems, while “taking a critical perspective” was defined as the tendency to question known conditions and asking for a chance of defense and the capability of assessing and evaluating reasons. The definitions of these two subscales are related to critical thinking. Thus, these two subscales were merged into the “taking a critical perspective” subscale. Lastly, the definition of “flexible thinking” is not very similar to other subscales. Thus, subscale 4 was not merged with any other subscale.

After the merging, there were five subscales in total. And the corresponding reliabilities were significantly higher between 0.528 and 0.797. In order to increase the correlations between the items, Items 8, 12, 16, 20, 23, 29, 35, 40, 42, 46, 47, 49, and 53 were removed based on the analysis results in Table 2. Moreover, because there were too many items in the “metacognition” subscale, Items 15, 36, and 55, which were the ones with rather low correlation coefficients, were removed. However, there were only six items in the “flexible thinking” subscale and the reliability of the subscale was 0.653, which was acceptable. Thus, none of the items in this subscale was removed.

In sum, a total of 16 items were removed. There were only 44 items left for the finalized version of the STDI (see Appendix). Then, reliability and validity analyses were performed to find out if the reliability and validity of this version were good.

The finalized version of the STDI and the reliability analysis. After the selection of subscales based on the item analysis results of the STDI draft, there were five subscales left, namely, persistence and managing impulsivity, thinking interdependently, thinking flexibly, metacognition, and taking a critical perspective (see Table 2 for the new item numbers after the revision and the original numbers). The contents of the subscales are summarized below.

There were 10 items in the “persistence and managing impulsivity” subscale—Items 1, 2, 10, 16, 17, 24, 25, 31, 37, and 38. The purpose of this subscale was to evaluate if the subject could work hard to hold on to the last moment without giving up, think before acting, keep his/her mind calm, ponder deeply over, and be careful and cautious. Examples of the items in this subscale include: Item 2 (When exploring a scientific question, I would not make any judgment until the right moment) and Item 10 (When encountering a difficult scientific

question that I cannot solve, I often just quit and give up).

There were eight items in the “thinking interdependently” subscale—Items 3, 8, 11, 15, 26, 30, 32, and 43. The purpose of this subscale was to evaluate if the subject could work together and truly cooperate with and learn from others in a mutually beneficial situation, try to understand others, put in some efforts to take others’ ideas and perspectives into consideration, and carefully appreciate others’ views and goodwill. Examples of the items in this subscale include: Item 8 (During a group discussion, I would comment on another person’s idea through a supportive way) and Item 15 (During a group discussion, I would provide positive opinions and opinions about improvements that can be made).

There were six items in the “thinking flexibly” subscale—Items 4, 12, 18, 27, 33, and 39. The purpose of this subscale was to evaluate if the subject could think from another perspective, find a way to change his/her views and innovate and come up with an alternative. Examples of the items in this subscale include: Item 18 (When exploring a scientific question, I would come up with different methods and solutions) and Item 27 (When exploring a scientific question, I would think from different perspectives).

There were 10 items in the “metacognition” subscale—Items 5, 6, 9, 19, 20, 23, 34, 36, 40, and 44. The purpose of this subscale was to evaluate if the subject could understanding his/her own thinking patterns, pay attention to his/her own thoughts, strategies, feelings, and conducts, think about the influences of these conducts on others, be capable of organized thinking, and solve problems with a plan. Examples of the items in this subscale include: Item 6 (When exploring a scientific question, I would review the goal to be achieved) and Item 20 (When exploring a scientific question, I would review the laws that should be followed).

There were 10 items in the “taking a critical perspective” subscale—Items 7, 13, 14, 21, 22, 28, 29, 35, 41, and 42. The purpose of this subscale was to evaluate if the subject could develop the attitude of seeking knowledge and asking questions, look for not only problems but also answers to these problems, and perform comparisons and make judgments based on the explanations, solutions, and evidences as well as their associations with the problems. Examples of the items in this subscale include: Item 14 (When exploring a scientific question, I would take the evidence against my viewpoint into consideration) and Item 35 (When exploring a scientific question, I would check the consistency between inferences).

Then, reliability analyses were performed, including internal consistency analyses and tests for associations among subscales. In the aspect of internal consistency, the α -value of the inventory was 0.952 and those of the five subscales—persistence and managing impulsivity, thinking interdependently, thinking flexibly, metacognition, and taking a critical perspective were 0.830, 0.778, 0.653, 0.866, and 0.836, respectively (see Table 4, in bold and italic font). Murphy and Davidshofer (2001) suggested that the α -value of a factor is considered good if above 0.70 and acceptable if above 0.60, while that of an inventory is considered good if above 0.80 and even better if above 0.90. All the α -value, except for the one of “thinking flexibly,” were above 0.70 and that of the overall inventory was above 0.90, meaning the reliability of this inventory was high.

In order to explore the correlations between the subscales, correlation analyses were performed. If the correlation coefficient of two variables is above 0.80, this correlation is considered strong. If it is between 0.40 and 0.80, this correlation is considered medium. If it is between 0.20 and 0.40, this correlation is considered weak. If it is under 0.20, there is almost no correlation at all (Murphy & Davidshofer, 2001).

As for the correlations between the subscales in this study, the obtained correlation coefficients were between 0.657 and 0.832 ($p < 0.000$) (see Table 4, in standard font). The correlations among “persistence and managing impulsivity,” “metacognition,” “taking a critical perspective,” and “metacognition” were strong, with

the coefficients of 0.823 and 0.832, respectively. And all the other between-subscale correlations were medium. However, because what these three subscales measured were different, it was not appropriate to merge these subscales. Thus, the 44 items were still arranged under five subscales.

Table 4

The α -Value of the Subscales and the Inventory After the Items to be Removed and the Correlations Among the Subscales

Subscale	Persistence and managing impulsivity	Thinking interdependently	Thinking flexibly	Metacognition	Taking a critical perspective
Persistence and managing impulsivity	0.830				
Thinking interdependently	0.693*	0.778			
Thinking flexibly	0.724*	0.657*	0.653		
Metacognition	0.823*	0.772*	0.739*	0.866	
Taking a critical perspective	0.756*	0.755*	0.728*	0.832*	0.836
Overall inventory	0.952				

Note. * $p < 0.001$.

The factor analysis of the STDI. Regarding factor analyses, Bandalos and Finney (2001) suggested to add up or group the item scores into parcel scores, and then, use these parcel scores as indicators, as the item idiosyncratic influence on parcel scores is rather small (Chapman & Tunmer, 1995). Moreover, the communality of parcel scores is usually rather high, which means that ratio of common-to-unique variance is rather large (Little, Cunningham, Shahar, & Widaman, 2002). Based on this argument, the researcher decided to randomly categorize the items of each subscale into two groups and added up the scores of the odd-numbered items and those of the even-numbered items (e.g., odd-numbered items under “metacognition” and even-numbered items under “metacognition”) for factor analyses.

The next step was to perform factor analyses. The obtained Kaiser-Meyer-Olkin (*KMO*)-value was 0.957 ($p < 0.001$), meaning that there were common factors among the items and it was suitable to perform factor analyses. Because the correlations between the subscales were strong or medium, the researcher applied principal axis factoring. The varimax method was adopted for oblique rotations. One factor—science thinking disposition was extracted, which was science thinking disposition. The factor loadings of the items were between 0.739 and 0.912 (see Table 5), 68.1% of the total variation of the 44 items inventory could be explained.

Table 5

The Factor Loadings of the STDI Subscales

	Science thinking disposition
Odd-numbered items under “metacognition”	0.912
Even-numbered items under “taking a critical perspective”	0.884
Even-numbered items under “metacognition”	0.876
Odd-numbered items under “persistence and managing impulsivity”	0.844
Odd-numbered items under “taking a critical perspective”	0.836
Even-numbered items under “thinking interdependently”	0.812
Even-numbered items under “persistence and managing impulsivity”	0.809
Even-numbered items under “thinking flexibly”	0.762
Odd-numbered items under “thinking interdependently”	0.762
Odd-numbered items under “thinking flexibly”	0.739

Why were not five factors extracted through the factor analysis? It was because that the item descriptions of the STDI were based on the theory proposed by Tishman, Jay, and Perkins (1993) and Costa and Kallick (2000). However, these scholars only built theoretical frames for thinking dispositions in their works. They did not verify the correlations between these factors. Thus, extracting only one factor—science thinking disposition, was also possible and reasonable. In addition, according to Table 4, the correlations between the subscales were strong or medium, meaning that when applying the STDI to research, explanations had to be made based on the total scores of the inventory. It was not appropriate to discuss any of the subscale scores individually.

Lastly, the structural equation model (SEM) was applied for confirmatory factor analysis (CFA). The validity of this inventory was examined based on the goodness-of-fit of the model. First of all, the measurement model plot was created based on the CFA results (see Table 5). Then, analysis of moment structures (AMOS), a statistical analysis software package, was used for analyses. Figure 1 was obtained and the ellipse in this figure was the one factor extracted, which was the latent variable. The five rectangles were the five manifest variables, namely, persistence and managing impulsivity, thinking interdependently, thinking flexibly, metacognition, and taking a critical perspective. The values above the upper-right corners of the rectangles were the R^2 -value from predicting the manifest variables using the latent variable and equaled to the communalities from general factor analyses. For example, the R^2 -value of “persistence and managing impulsivity” was 0.75 with e_1 - e_5 being the measurement errors. Thus, the “persistence and managing impulsivity” = thinking disposition \times 0.87 + e_1 . The “thinking interdependently” = thinking disposition \times 0.82 + e_2 .

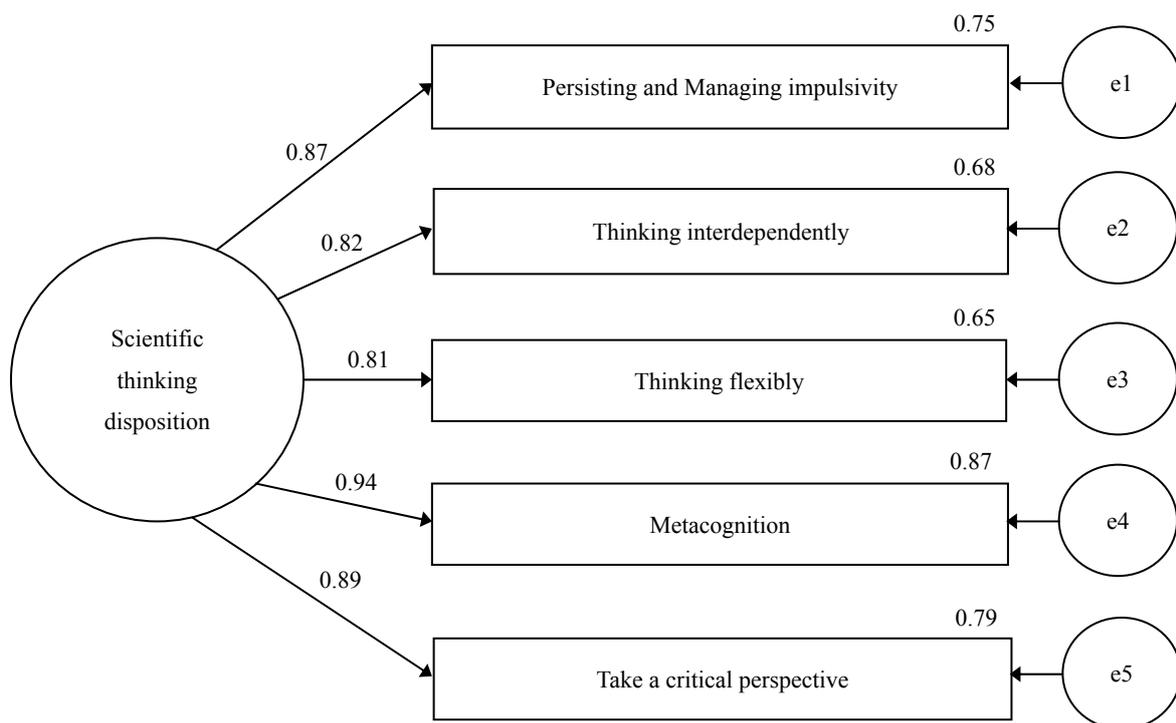


Figure 1. The measurement model of the STDI (standardized solution).

As for the goodness-of-fit test for the measurement model, the maximum likelihood (ML) method was applied for parameter estimation. In the aspect of preliminary fit criteria, all the error variances were positive between 4.612 and 30.397. All significant ($p < 0.001$) with the factor loadings between 0.81 and 0.94 (see

Figure 1), which was between 0.5 and 0.95, and the absolute values of correlation coefficients between parameter estimates between 0.003 and 0.609, which was not too close to 1. These results showed that the model being tested was good and the latent variables could be measured by using the observed variables.

Table 6

Summary of the Overall Goodness-of-Fit Indicators of the STDI

Indicator	χ^2	<i>df</i>	<i>p</i>	χ^2 ratio	GFI	AGFI	IFI	NFI	SRMR
Value	11.381	5	0.044	2.276	0.986	0.957	0.995	0.991	0.0125

Then, the overall goodness-of-fit of the model was tested. Table 6 shows that $\chi^2 = 11.381$ ($df = 5$). And statistical significance was reached. However, due to the common influence of sample size on χ^2 , χ^2 was not considered as a good indicator for goodness-of-fit of the model. In addition, the χ^2 ratio was 2.276 (lower than 3), the acceptance threshold. And indicators, such as goodness-of-fit index (GFI), adjusted goodness-of-fit index (AGFI), incremental fit index (IFI), normal fit index (NFI), and standardized root mean square residual (SRMR) were less likely to be influenced by sample size. Thus, they were used to evaluate the goodness-of-fit of the model. It was found that SRMR = 0.0125 (lower than 0.05), while GFI, AGFI, IFI, and NFI were all above 0.9, showing that the overall goodness-of-fit was good, and the hypothesis could be accepted (Arbuckle & Wothke, 1999). In other words, the measurement model of the inventory fitted the actual data and there was evidence to support the test model proposed by the researcher.

The fit of the internal structure of the model was tested (see Table 7). In the aspect of individual subscale reliability, the α -value of “persistence and managing impulsivity,” “thinking interdependently,” “thinking flexibly,” “metacognition,” and “taking a critical perspective” were 0.754, 0.68, 0.649, 0.874, and 0.794, respectively (all above 0.5). The latent variable reliability was 0.938 (above 0.6). All the parameter estimates were significant ($p < 0.001$). The latent variable average variance extracted (AVE) was 0.750 (above 0.5). The absolute values of standardized residuals were between 0.021 and 0.348 (under 1.96). The modification indices were between 0.000 and 2.740 (under 3.84). Thus, the fit of the internal structure of the model was good.

Table 7

Summary of the Indices for the Fit of the Internal Structure of the Model for the STDI

Index	Reliability		Latent variable AVE	Parameter estimate significance	Absolute value of standardized residual	Modification index
	Individual subscale	Latent variable				
Result	0.648-0.874	0.938	0.750	All under 0.001	0.021-0.348	0.000-2.740

After integrating the test results of the preliminary fit criteria, overall model fit and fit of internal structure, it was found that the STDI met all the standards, meaning it was a good measurement tool.

Stage II

The construct validity of the STDI. The literature review above shows that scientific thinking dispositions can significantly influence scientific creativity and scientific reasoning. Thus, the researcher inferred that scientific thinking dispositions could be relate to scientific creativity and scientific reasoning. Those with high STDI scores would score high in the scientific creativity test and the scientific reasoning test, while those with low STDI scores would score low in the scientific creativity test and the scientific reasoning test. Based on this hypothesis, the researcher used the differences in the performances of the scientific creativity

test and the scientific reasoning test between the student groups of different scientific thinking dispositions as the proof for the construct validity of the STDI. Thus, in Stage II, the tests were performed with the research sample.

Then, the correlations between scientific thinking dispositions and scientific creativity and between scientific thinking dispositions and scientific reasoning were examined (see Table 8). According to the results, the correlation between scientific thinking dispositions and scientific creativity was 0.288 ($p = 0.000$), and that between scientific thinking dispositions and scientific reasoning was 0.337 ($p = 0.000$), which are both low and positive correlations. Although these correlations were low, they were still statistically significant. In other words, they were robust and meaningful (Fraenkel & Wallen, 2006). That means, theoretically, there are correlations between scientific thinking dispositions and scientific creativity, and correlations between scientific thinking dispositions and scientific reasoning, and this study found evidences for that.

Table 8

Correlations Between Scientific Thinking Dispositions and Scientific Creativity and Correlations Between Scientific Thinking Dispositions and Scientific Reasoning

	Scientific thinking dispositions total score	Scientific creativity total score
Scientific creativity total score	0.288*	
Scientific reasoning total score	0.337*	0.541*

Note. * $p < 0.001$.

Based on these findings, further tests were conducted to examine the differences in performances of the scientific creativity test and the scientific reasoning test between the student groups of different scientific thinking dispositions, for the purpose of exploring the construct validity of the STDI.

The differences in performances of the scientific creativity test between the high-STDI-score group and the low-STDI-score group. First of all, the students were ranked based on their STDI scores. The top 27% were assigned to the high-STDI-score group, while the bottom 27% to the low-STDI-score group. Table 9 shows that there were 46 students in the high-STDI-score group with the mean score (*Mean*) and standard deviation (*SD*) being 169.13 and 13.633, respectively, and there were 47 students in the low-STDI-score group, with the *Mean* and *SD* being 110.62 and 16.519, respectively.

Table 9

The Score Performances of the High-STDI-Score Group and the Low-STDI-Score Group

Thinking disposition group	<i>N</i>	<i>Mean</i>	<i>SD</i>
High-STDI-score group	46	169.13	13.633
Low-STDI-score group	47	110.62	16.519

Then, the independent sample *t*-test was performed on the scientific creativity test results. The Levene value of the homogeneity of variance test of the high-STDI-score group and the low-STDI-score group was significant ($F = 10.283$, $p = 0.002$), meaning the two samples showed significant difference in the dispersion of scientific creativity. Thus, adjusted *t*-test was more appropriate (Aron & Aron, 2003; Howell, 2007).

Table 10 shows the results of the independent sample *t*-test, with $t = 3.739$ and $p = 0.000$, meaning that the performances of the scientific creativity test between the high-STDI-score group and the low-STDI-score group were different. That is, the students with high STDI scores would score high in the scientific creativity test.

Table 10

The Scientific Creativity Test Scores of the High-STDl-Score Group and the Low-STDl-Score Group and the T-Test Results

Subscale	Scientific thinking disposition group	<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>t</i>	<i>Sig.</i>
Scientific creativity	High-STDl-score group	46	39.04	23.602	3.739	0.000*
	Low-STDl-score group	47	24.21	13.047		

Note. * $p < 0.001$.

The differences in performances of the scientific reasoning test between the high-STDl-score group and the low-STDl-score group. Then, the test was performed on the scientific reasoning test results. The Levene value of the homogeneity of variance test of the high-STDl-score group and the low-STDl-score group was significant ($F = 4.964$, $p = 0.028$), meaning the two samples showed significant difference in the dispersion of scientific reasoning. Thus, adjusted t -test was more appropriate (Aron & Aron, 2003; Howell, 2007).

Table 11 shows the results of the independent sample t -test, with $t = 3.808$ and $p = 0.000$, meaning that the performances of the scientific reasoning test between the high-STDl-score group and the low-STDl-score group were different. That is, the students with high STDl scores would score high in the scientific reasoning test.

Table 11

The Scientific Reasoning Test Scores of the High-STDl-Score Group and the Low-STDl-Score Group and the T-Test Results

Subscale	Scientific thinking disposition group	<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>t</i>	<i>Sig.</i>
Scientific reasoning	High-STDl-score group	46	5.48	3.244	3.808	0.000*
	Low-STDl-score group	47	3.11	2.737		

Note. * $p < 0.001$.

In sum, the students with high STDl scores would score high in the scientific creativity test and the scientific reasoning test. Moreover, it was found in Stage I that the items of the STDl all met the criteria to contrast an inventory. This is sufficient to show that the STDl was a good measurement tool.

Conclusions and Implications

This study aimed to develop a STDl with good reliability and validity. Through the processes of literature review to explore the implications of scientific thinking dispositions, developing the draft of the STDl, performing a pilot study and item analyses, and removing items based on the item analysis results, the STDl was created with 44 items, covering contents in relation to persistence and managing impulsivity, thinking interdependently, thinking flexibly, metacognition, and taking a critical perspective. Firstly, the result of the factor analysis showed that only one factor—scientific thinking disposition, could be extracted, meaning that when applying the STDl to research, the results must be explained based on the total scores and it is not appropriate to discuss any of the subscales separately. Secondly, the obtained Cronbach's α of the STDl was 0.952, meaning that the reliability of this inventory was good. Moreover, the measurement model of the inventory met the fit indices, meaning that the STDl was a good measurement tool. In addition, the differences in the performances of the scientific creativity test between the high-STDl-score group and the low-STDl-score group was significant ($p < 0.001$), meaning the construct validity of the inventory was good. All these

evidences showed that the results of the item analyses and the validity and reliability tests were ideal, so the STDI was an efficient tool for assessing eighth graders' scientific thinking dispositions.

Based on the conclusions above, this study proposed three suggestions as below:

1. Developing scientific thinking dispositions is important for students when it comes to learning science (AAAS, 1989). Thus, future studies can adopt the STDI developed by this study to explore the differences in concept learning, problem solving, scientific argument, and inquiry skills between students of different scientific thinking dispositions. This way, the construct validity of this inventory can be further supported by more proofs. And the possible associations of students' scientific thinking dispositions with their various performances in learning science can be further investigated.

2. By using this inventory with other scientific thinking related variable for research, the causal model of influences on students' ways of thinking can be further studied, and the obtained results may help with the overall understanding of students' scientific thinking.

3. The research results showed that there were differences in the performances of the scientific creativity test between student groups of different scientific thinking dispositions, meaning that scientific thinking dispositions are important for scientific creativity and scientific reasoning. Therefore, teachers can use this inventory to understand their students' scientific thinking dispositions, to examine their changes in scientific thinking dispositions to further explore their learning characteristics, or to design teaching activities that can help students to develop scientific creativity and scientific reasoning capabilities.

References

- Alsop, S., & Watts, M. (2003). Science education and affect. *International Journal of Science Education*, 25(9), 1043-1047.
- American Association for the Advancement of Science (AAAS). (1989). *Science for all Americans*. New York, N.Y.: Oxford University Press.
- Arbuckle, J., & Wothke, W. (1999). *AMOS 4.0 user's guide*. Chicago: Smallwaters Corporation, Inc..
- Aron, A., & Aron, E. N. (2003). *Statistics for psychology* (3rd ed.). New Jersey: Prentice Hall.
- Bandalos, D. L., & Finney, S. J. (2001). Item parceling issues in structural equation modeling. In G. A. Marcoulides (Ed.), *New developments and techniques in structural equation modeling* (pp. 269-296). Mahwah, N.J.: Lawrence Erlbaum Associates Publishers.
- Bransford, J., Brown, A. M., & Cocking, R. (1999). How people learn: Brain, mind, experience, and school. In *Report of the National Research Council*. Washington, D.C.: National Academy Press.
- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116-131.
- Chapman, J. W., & Tunmer, W. E. (1995). Development of young children's reading self-concepts: An examination of emerging sub-components and their relationship with reading achievement. *Journal of Educational Psychology*, 87(1), 154-167.
- Chen, C. M. (2004). A study of the factors affecting the scientific creativity of 10th grade students (Unpublished doctoral dissertation, National Kaohsiung Normal University).
- Clarke, T. E. (2002). Unique features of an R & D work environment and research scientists and engineers. *Knowledge, Technology & Policy*, 15(3), 58-69.
- Costa, A., & Kallick, B. (2000). *Discovering and exploring habits of mind*. Alexandria, V.A.: Association for Supervision of Curriculum Development.
- Fraenkel, J. R., & Wallen, N. E. (2006). *How to design and evaluate research in education* (6th ed.). New York, N.Y.: McGraw-Hill.
- Howell, D. C. (2007). *Statistical methods for psychology* (6th ed.). Belmont: Thomson Wadsworth.
- Hu, W., & Adey, P. (2002). A scientific creativity test for secondary school students. *International Journal of Science Education*, 24(4), 389-404.
- Lawson, A. E. (1995). *Science teaching and the development of thinking*. Belmont, C.A.: Wadsworth.
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, 9(2), 151-173.

- Louis, K. S., Holdsworth, J. M., Anderson, M. S., & Campbell, E. G. (2007). Becoming a scientist: The effects of work-group size and organizational climate. *Journal of Higher Education, 78*(3), 311-336.
- McGrayne, S. B. (1998). *Nobel Prize women in science* (2nd ed.). Washington, D.C.: Joseph Henry Press.
- McIntire, S. A., & Miller, L. A. (2007). *Foundations of psychological testing: A practical approach*. Thousand Oaks: Sage.
- Murphy, K. R., & Davidshofer, C. O. (2001). *Psychological testing: Principles and applications* (5th ed.). Sydney: Pearson Education Australia.
- National Research Council (NRC). (1996). *National science education standards*. Washington, D.C.: National Academy Press.
- Patrick, H., & Yoon, C. (2004). Early adolescents' motivation during science investigation. *Journal of Educational Research, 97*(6), 319-328.
- Perkins, D. N., & Tishman, S. (1998). *Dispositional aspects of intelligence*. Retrieved from <http://learnweb.harvard.edu/alps/thinking/docs/Plymouth.pdf>
- Perkins, D., Tishman, S., Ritchhart, R., Donis, K., & Andrade, A. (2000). Intelligence in the wild: A dispositional view of intellectual traits. *Educational Psychology Review, 12*(3), 269-293.
- Settlage, J., & Southerland, S. A. (2007). *Teaching science to every child*. New York, N.Y.: Routledge.
- Shieh, F. P., & Hung, J. F. (2008, February). Thinking dispositional view on scientific reasoning and scientific creativity performance. In *Conference of Asian Science Education (CASE)*. Taiwan: Kaohsiung.
- Siegel, H. (1999). What (good) are thinking dispositions? *Educational Theory, 49*(2), 207-222.
- Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology, 89*(2), 342-357.
- Stanovich, K. E. (2001). The rationality of educating for wisdom. *Educational Psychologist, 36*(4), 247-251.
- Tishman, S., Jay, E., & Perkins, D. N. (1993). Teaching thinking dispositions: From transmission to enculturation. *Theory Into Practice, 32*(3), 147-153.
- Tuan, H. L., Chin, C. C., & Shieh, S. H. (2005). The development of a questionnaire to measure students' motivation towards science learning. *International Journal of Science Education, 27*(6), 639-654.

Appendix: Modified Item Descriptions for the STDI

Item No.	Description
1	When learning science, no matter how difficult the problem I encounter is, I would persist in achieving my goal and I would never give up easily.
2	When exploring a scientific question, I would not make any judgment until the right moment.
3	When discussing a scientific question with others, I would stop my own thinking for a moment to carefully consider others' perspectives.
4	When exploring a scientific question, I would come up with various ideas based on data from multiple sources.
5	When exploring a scientific question, I would monitor my own thought to find my bias.
6	When exploring a scientific question, I would review the goal to be achieved.
7	When learning science, I often think, "If ..., what would happen?"
8	During a group discussion, I would comment on another person's idea through a supportive way.
9	When exploring a scientific question, I would organize various ideas in a logical way.
10	When encountering a difficult scientific question that I cannot solve, I often just quit and give up.
11	When discussing a scientific question with others, I can learn diversified perspectives from others.
12	When exploring a scientific question, I do not know what to do if my first solution does not work.
13	When learning science, I often think, "What is the connection between these things?"
14	When exploring a scientific question, I would take the evidence against my viewpoint into consideration.
15	During a group discussion, I would provide positive opinions and opinions about improvements that can be made.
16	When exploring a scientific question, even in a chaotic and confusing situation, I would not give up and would hold on till the last moment.
17	When exploring a scientific question, I would patiently confirm the answer in my mind before giving it.
18	When exploring a scientific question, I would come up with different methods and solutions.
19	When exploring a scientific question, I would think about strategies for future improvement after the exploration is completed.
20	When exploring a scientific question, I would review the laws that should be followed.
21	When learning science, I often think, "What is the cause of this relationship?"
22	When exploring a scientific question, I would compare the advantages and disadvantages between different solutions.
23	When exploring a scientific question, I would divide the question into several parts in an organized way.
24	When exploring a scientific question, I would not back down no matter how many frustrations come along.
25	When exploring a scientific question, I would spend some time assessing other alternatives and possible results.
26	When discussing a scientific question with others, I would extend their perspectives and clarify with them.
27	When exploring a scientific question, I would think from different perspectives.
28	When exploring a scientific question, I would also look for issues that can be further looked into after making a conclusion.
29	When exploring a scientific question, I am not good at determining which information is related to the question and which is not.
30	During a group discussion, I would discuss with others about different views.
31	When I encounter a scientific question that I cannot solve, I would just come up with an answer and get it over with.
32	When discussing a scientific question with others, I would interpret their perspectives accurately.
33	When exploring a scientific question, I could accept reasoning that is completely opposite to my viewpoint.
34	When exploring a scientific question, I would keep on reviewing my strategies to make sure they are appropriate.
35	When exploring a scientific question, I would check the consistency between inferences.
36	When exploring a scientific question, I would do it step by step and logically.
37	If I cannot complete the exploration of a scientific question when I try for the first time, I would keep on trying until I succeed.

38	When exploring a scientific question, I would conduct experiments over and over until I am satisfied with the results before writing down my conclusion.
39	When exploring a scientific question, I am not good at thinking about the question from different perspectives.
40	When learning science, I would look over my assignment again after it is completed to try to correct all mistakes if possible without being reminded.
41	When learning science, I would ask questions actively in order to obtain new information.
42	When exploring a scientific question, I would make sure my conclusion is based on current evidences and reasonable reasons.
43	During a group discussion, other people's views often help me to come up with a new question or a new thinking direction.
44	I would explore scientific questions with plans and methods.