# Cognitive and Attentional Process in Insight Problem Solving of the puzzle game "Tangram"

**Yoshiki Nakano (nakano@ed.akita-u.ac.jp)**
Faculty of Education and Human Studies, Akita University. Tegata Gakuen-machi 1-1.
Akita City, 010-8502 JAPAN

**Abstract**

The purpose of this study is to demonstrate a constraint relaxation which is followed by the transition to an appropriate representation in insight problem solving. The puzzle game "Tangram" was used as a new insight problem, in which problem-solvers were presented a silhouette and asked to make the same configuration by arranging 7 pieces. At the beginning, problem-solvers had a constraint allocating the pieces into a geometric shape, but then relaxed this to reach the correct configuration at a later stage of problem solving. Participants' subjective assessments of their confidence to reach the solution predicted neither the constraint relaxation nor the successful problem solving. However, eye-tracking data suggested that the successful problem-solvers tended to search the problem space more widely than the unsuccessful-problem solvers.

**Keywords:** Tangram; insight problem solving; constraint relaxation; eye tracking.

## Introduction

The Insight problem solving, in general, exhibits a characteristic pattern as a following process (Kaplan & Simon, 1990; Metcalfe, 1986). At first, an insight problem seems easy to solve, but problem-solvers are caught in an impasse soon after. They get stuck, think that all options have been explored and lose their way. When a sudden and useful idea comes to mind, it often leads problem-solvers rapidly to the solution.

There is an agreement that inappropriate constraints for a solution are the main source of the difficulty to solve an insight problem (Jones, 2003; Orrmerod, MacGregor, & Chronicle, 2002). These studies suggested that insight requires the relaxation of such inappropriate constraints, and that an impasse can be broken by changing a representation of the problem. A constraint is a tendency of thinking and behavior that is taken in attempting to solve a certain problem. It usually facilitates the process to reach the solution. When a self-imposed constraint, however, is inappropriate to solve a problem, it prompts a critical difficulty to achieve insight, as it activates irrelevant knowledge and causes attempts that cannot contribute to correct solution. This leads problem-solvers into an impasse. In spite of the consensus about the source of difficulty of insight problems, a dynamic process to reach a solution has not been identified clearly. The purpose of this study is to specify an inappropriate constraint which inhibited the insight into the solution in a geometric problem solving, and to provide direct evidence about a critical factor for the successful problem solving. In addition, an eye-tracking technique is adopted to examine whether a proper searching of the problem space can lead successful problem solving or not (Thomas & Lleras, 2009).

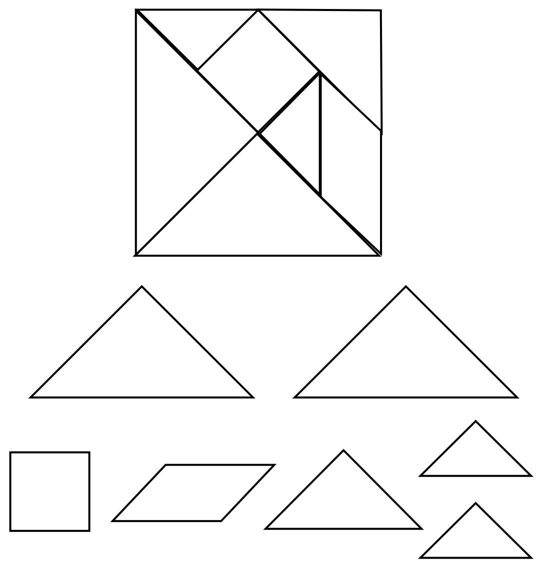


Figure 1: The 7 pieces of Tangram. Top: a configuration when the pieces are divided from a square. Bottom: Separated seven pieces. A problem-solver of Tangram move, rotate or combine them to complete a task silhouette.

This study also aims to test whether the difficulty of the constraint relaxation can be reflected in an apparent task performance of a problem-solver while independent from subjective awareness. For these purposes, the puzzle game of "Tangram" was used as a new insight problem. Tangram consists of 7 triangular, square or parallelogram pieces (Figure 1). Problem-solvers are presented a task silhouette and required to make the same configuration by arranging these pieces. In Tangram, numerous task silhouettes can be composed by the 7 pieces; for example geometrical shapes, animals or objects. Because each silhouette has an individual configuration of the pieces, problem-solvers often cannot find a correct configuration immediately after the silhouette has been presented. They usually repeat trials with arranging the pieces until the solution is completed.

Tangram has the advantages of allowing researchers to monitor task performance of a problem solving requiring insight. Nakano (2009) recorded protocols of participants and movements of the pieces. His protocol data revealed that the participants who could complete the task silhouette had expressed an "Aha" experience before reaching the correct configuration. Nakano (2009) found that the participants tended to combine the pieces into a geometric shape such as a square or a triangle. This tendency was usually involved combinations of the 2 largest triangles among the 7 pieces. Such the constraints will facilitate the process to reach the solution, in the case that a correct configuration includes geometric combination of the pieces as a square or a triangle. For example, the task silhouette of an arrow has two patterns of the correct configuration, and the both include geometric combination of the 2 largest triangles (bottom in Figure 2). However, when these 2 pieces must be combined into an irregular configuration to complete a task silhouette, the tendency to construct a geometric shape will inhibits the insight into the correct configuration. For example, to complete the silhouette of a lion (top-left in Figure 2) the 2 triangle pieces must be attached by sliding their longer sides in an opposite direction to each other (top-right in Figure 2). Most problem-solvers who tried to complete this silhouette reported that it was difficult to discover this irregular configuration (Nakano, 2009). This finding indicated that the problem-solvers imposed apparent constraints on allocations and combinations of the pieces in the problem solving of Tangram.

The primary purpose of the present experiment is to investigate the correspondence between the constraint relaxation and explicit movements of the pieces. For this purpose, movements of the 7 pieces were recorded and combinations of the 2 largest triangle pieces were identified. These combinations will change as the initial constraint is relaxed and more appropriate representations are constructed. Participants who try to complete the task silhouette of an arrow will not have to relax the constraint to allocate the 2 largest triangle pieces in a geometric pattern. So they will reach the solution easier and faster comparing to the task silhouette of a lion. When they try to complete this task silhouette, they find it impossible to complete the solution only by arranging the pieces into familiar geometric combinations. Then, the participants arrange the pieces into other combinations. To investigate this hypothesis, arrangements of the 2 largest triangle pieces are categorized into three categories, geometric, transitional or irregular combinations. The geometric means the combination of the 2 pieces was a square or a triangle. The irregular combination is that the 2 pieces have a contact each other but no corners of the pieces meet or the edges of the pieces were placed adjacently without combining their corners. The rests of combinations, as that the pieces were arrange in the same direction or in a symmetric configuration, are categorized in the transitional combination.

The second purpose is to examine independence between the constraint relaxation and subjective awareness in the problem solving of Tangram. In Nakano (2009), participants were required to evaluate how confident they could complete the correct configuration before and during problem solving of Tangram. The subjective confidence decreased over time even in the successful problem-solvers. Thus, the subjective evaluation did not predict performance on the insight problem, as had been suggested by Metcalfe and Wiebe (1987). This finding supports the idea that the transition from one rule to a more appropriate representation for a solution proceeds without subjective awareness.

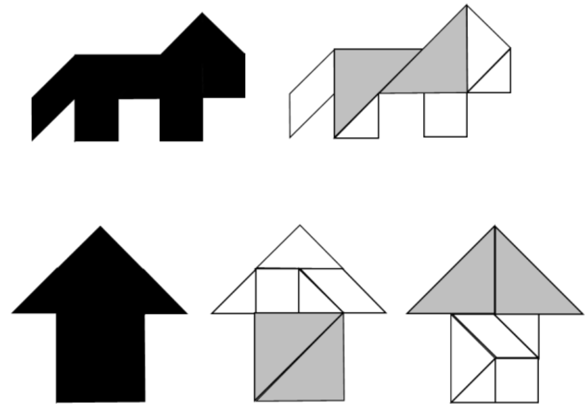


Figure 2: Top-left; the task silhouette of a lion. Top-right; the correct configuration of the 7 pieces constructing the silhouette of a lion. The two largest triangle pieces (greyed part) are allocated in irregular combination. Bttom-left; the task silhouette of an arrow. Bottom-center and -right; the correct configuration constructing the silhouette of an arrow. There are two patterns of correct configuration. In both, the 2 pieces are allocated in a geometric combination.

## Methods

### Participants

Twenty undergraduates (mean age = 20.8, age range = 20-26) participated in the experiment. All were naive to Tangram.

### Apparatus

The Tangram was comprised of 7 pieces. A set of the pieces was made by dividing a square plate which was 11.8cm in width and in height (Figure 1). The task silhouette was a lion or an arrow (Figure 2 left). The size of the both silhouettes were the same as one-fifth of the correct configuration which constructed by the 7 pieces. During the task, participants assessed how confident they were to complete the task at the Graphics Rating Scale (GRS). The GRS involved 6 verbal descriptors and 13 scale marks along a horizontal line. A description "Not at all" was written at the left end of the horizontal line, and "Nearly completed" was at the right end. An index arrow was attached on the GRS board so that the participants could indicate their positions on the horizontal line taking the descriptions into account. At the beginning of the experiment, the index arrow was located on the center of the 13 scale marks. A digital video camera (Panasonic NV-GS100) recorded movements of the pieces and a location of the index arrow from 65 cm above.

### Procedure.

A participant sat down in front of a work desk on which the 7 pieces of Tangram and the GRS board were situated. The pieces were allocated on the desk as depicted in Figure 1 (below). After the instructions about what the experiment involved, the participant was presented the task silhouette which was printed in black and was required to complete the configuration of it by using all the pieces. The participant indicated how he or she would be able to complete the correct configuration of the silhouette by the index arrow on the GRS board. During the assessment, the participant was allowed to see the pieces and the silhouette but was not permitted to touch or move the pieces. After the assessment, the participant was allowed to move the pieces. Each session lasted for 240 sec, and was followed by the assessment on GRS. Then 1 minute of rest period was given to the participant until the next session was started.

When a participant completed the correct configuration, the problem solving was successfully ended. A sum of the manipulation time until the completion was accounted as a solution time of the "completer". A participant who could not complete the correct configuration until the end of session 5 was counted as a "non-completer". In this case, the total manipulation time was 20 min.

All the participants participated in the two-day experiment. Either the silhouette of the lion (lion task) or the arrow (arrow task) was given to the participant as a task in each day. The order of the task silhouettes was counterbalanced among the participants.

### Recording of eye movements

During process of the problem solving of Tangram, eye movements were captured using Talk Eye Lite (Takei Scientific Instruments Co. Ltd, Japan). Eye tracking was operated in monocular mode on the right eye and at a sampling rate of 33 Hz. The participants were seated on a chair and their head was fixed by using a chin and forehead rest to keep a distance from a surface of a work desk to eye in approximately 40 cm. The surface of the work desk was tilted to about 10 degree angle. The 7 pieces of Tangram were located on the desk surface and were moved within the range of 25 cm in length and 35 cm in width so as not to go outside the participant's eye-field.

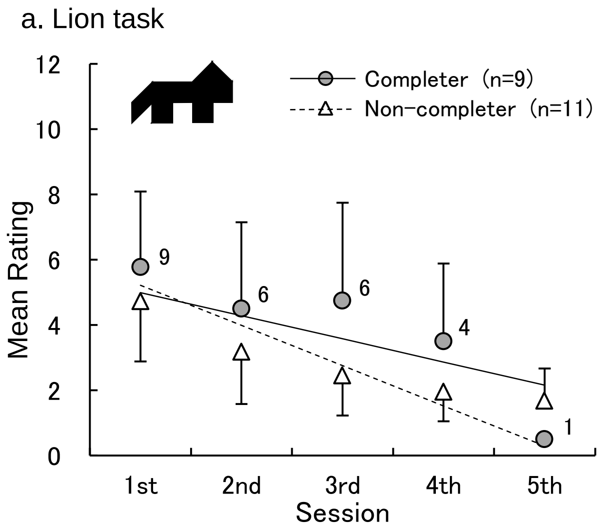
## Results

Among the 20 participants, 9 completed the correct configuration of the task silhouette of a lion. The mean solution time was 594.7 sec (*SD* = 376.0 sec, *Min* = 106 sec, *Max* = 1080 sec). Three of the 9 completers had finished in session 1, and 2 of the remaining 6 had finished in session 3. Three had finished in session 4 and 1 had finished in the last session. In the arrow task, 12 participants completed the correct configuration within 5 sessions. The mean solution time was 483.6 sec (*SD* = 331.1 sec, *Min* = 58 sec, *Max* = 1030 sec). Among the 12 participants 4 had finished in session 1. Two of the remaining 8 completers had finished in session 2. Three of the remaining 6 completers had finished in session 3. Two had finished in session 4 and 1 had finished in the last session. Eight of the 20 participants completed both the lion and the arrow task. The averaged solution time of these "high-achievers" was faster in the arrow task (532.8 sec) than in the lion task (644.8 sec), but statistical analysis revealed that the difference did not reach significant level ($p > .1$). Therefore, contradicting to the expectation, both the ratios of the completer and their solution time did not indicate that there is a difference in difficulty between the two task silhouettes.

Ratings of subjective confidence to the completion were identified from a location of the index arrow on GRS from the video image. The value of 0 was assigned to the mark on left extreme side, and 12 was assigned to the right extreme mark. Six assigned to the central mark which located between the descriptions "possibly, I can complete" and "possibly, I cannot complete". A value of each rating was identified on the basis of relative distance of the index arrow from the marks. Mean ratings were calculated among the completers and the non-completers in each sessions. Figure 3 show the plots of the mean ratings as a function of the sessions for each task silhouette. Because the participants who had completed the task silhouette were not engaged in the assessment anymore, sample numbers included in the plots of the completer decreased as the session proceeded. All mean ratings of non-completers included 11 participants in Figure 3a and 8 in Figure 3b.

To investigate the relationship between participants' confidence and results of the problem solving, they were categorized into a high-achiever who completed the both task silhouettes or into a non-achiever who could not complete the both silhouettes. Thus, 5 participants who only completed the either task were eliminated from this analysis. Table 1 shows mean ratings of confidence in session 1 and in the last session for the high-achievers ($n$ = 8), excepting one participant who completed the both tasks in session 1, and for the non-achievers ($n$ = 7), respectively. For the completers, the last session is when they completed the task, and for the non-completers it was always the fifth session. A $2 \times 2$ analysis of variance (ANOVA) with the variables achievement (high-achiever vs. non-achiever) and time (session vs. last session) showed that the effect of time was significant, $F(1, 12) = 52.2$, $p < .01$. In detail, the mean rating in session 1 was higher than that in the last session in the both groups. Supported with this result, the linear regression lines in Figure 3 indicated that participants' subjective assessments of their confidence decreased over time. The main effect of achievement and the interaction between the two variables did not reach statistical significance.

To investigate whether constraint relaxation was reflected on actual manipulation of the pieces, arrangements of the 2



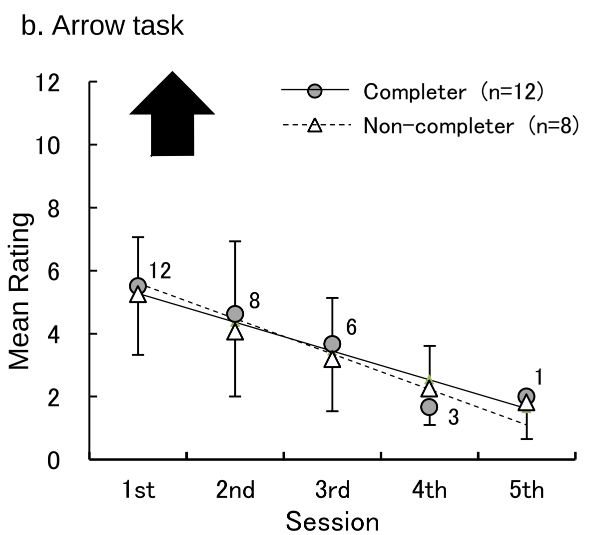




Figure 3: Plots of mean ratings of subjective confidence and linear regression lines in the lion task (a) and the arrow task (b). The plots of Grey circle with a solid line indicate results of the completers and the plots of blank triangle with a dashed line indicate results of the non-completers. The numbers beside the plots of the completer indicate a number of samples included in each mean value. Error bars indicate *SD*s.

Table 1: Mean ratings and SDs of a participants' confidence

| | High-achiever | Non-achiever |
|---|---|---|
| 1st session | 5.79 (1.8) | 4.82 (1.6) |
| Last session | 3.36 (2.6) | 1.79 (0.8) |

largest triangle pieces were categorized and their duration times were compared. When these 2 triangle pieces were in contact with each other at least on a corner or a side of them, this arrangement was counted as a combination of the pieces. All combinations of the 2 largest triangle pieces were identified on the video image and categorized into a geometric, a transitional or an irregular combination. In the case that the 2 pieces were combined into a triangle or a square, such the allocation was categorized into the geometric combination. In the arrow task, this category includes the correct combination of the 2 pieces. The participants often allocated the 2 pieces such that the short edge of one triangle is placed adjacent to the long edge of the second triangle so that no corners of the sides met, additionally the corner of the long edge of the second triangle did not make contact with the first triangle. As with these examples, an allocation of the 2 pieces in which the edges were placed adjacently without combining their corners was categorized into the irregular combination. In the lion task, this category includes the correct combination of the 2 pieces. All other combinations that were categorized neither into the geometric nor into the irregular combination were categorized into the transitional combination. A measure of the duration of the time in which each combination was made was taken from the video image. The measurement started from the moment at which the 2 pieces were allocated into a certain combination and ended when they were separated or changed into the other arrangement. To investigate a transition of a predominant combination of the pieces, overall manipulation time of each participant was divided into a first and a second half period. A cumulative duration time of each combination was calculated for each participant. Table 2 shows mean percentage of the cumulative duration time of each combination over the total manipulation time.

Three-way ANOVA with the variables completion (completer vs. non-completer), time ($1^{st}$ vs. $2^{nd}$ half) and combination (geometric, transitional or irregular) was computed on percentage of the cumulative duration time. In the lion task, the results showed that there was a significant main effect of time variable, $F(1, 18) = 5.1$, $p < .05$, and two-way interaction between completion and time, $F(1, 18) = 7.9$, $p < .05$. Post hoc analysis revealed that, in the completers, an averaged percentage of the cumulative duration time over the three combinations was significantly higher in the second half (14.9%) than in the first half (6.0%), $F(1, 18) = 12.7$, $p < .01$, but the difference was not significant in the non-completer (11.3% in the $1^{st}$ half; 10.3% in the $2^{nd}$ half). Additionally, two-way interaction between time and combination was also significant, $F(2, 36) = 7.5$, $p < .01$. Post hoc analysis revealed that the percentage significantly increased than the first half (5.7% in the $1^{st}$ half vs. 17.9% in the $2^{nd}$ half), $F(1, 54) = 12.8$, $p < .01$, while for the transitional combination the increase was marginally significant (8.2% vs. 14.8%), $F(1, 54) = 3.7$, $p < .06$. In contrast, for the geometric combination, the percentage in the first half was significantly higher than the second half

Table 2: Percentage of the cumulative duration time of each combination over the total manipulation time

| Lion task | Completer | Non-completer |
|---|---|---|
| Geometric | | |
| 1st half | 9.1% | 15.2% |
| 2nd half | 2.3% | 8.1% |
| Transitional | | |
| 1st half | 6.0% | 10.4% |
| 2nd half | 13.1% | 16.5% |
| Irregular | | |
| 1st half | 3.1% | 8.3% |
| 2nd half | 29.3% | 6.3% |
| **Arrow task** | **Completer** | **Non-completer** |
| Geometric | | |
| 1st half | 29.0% | 19.2% |
| 2nd half | 26.8% | 27.1% |
| Transitional | | |
| 1st half | 1.4% | 1.7% |
| 2nd half | 6.1% | 1.5% |
| Irregular | | |
| 1st half | 5.3% | 4.9% |
| 2nd half | 5.3% | 3.7% |

(12.1% vs. 5.2%), $F(1, 54) = 4.1$, $p < .05$. The three-way interaction reached statistically significant, $F(2, 36) = 4.9$, $p < .05$. Post hoc analysis revealed that in the second half the percentage of the irregular combination was higher in the completer than that in the non-completer (29.3% vs. 6.3%), $F(1, 54) =13.8$, $p < .01$, but in the first half the difference was not significant (3.1% vs. 8.3%). Regarding the completers, the percentage of the irregular combination (29.3%) was significantly higher than that of the geometric (2.3%) and the transitional combination (13.1%) in the second half, all *p*s < .01. This percentage of the completer in the second half was significantly increased than the first half (3.1%), $F(1, 54) =29.8$, $p < .01$. In the arrow task, three-way ANOVA showed that the main effect of combination variable was significant, $F(2, 36) = 11.1$, $p < .01$. Multiple comparisons revealed that the percentage of the geometric combination averaged over time and completion variables (25.2%) was significantly higher than that of the transitional (2.8%) and the irregular combination (5.1%), all *p*s < .01. The other main effect and the interactions did not reach statistical significance.

To investigate whether there is a characteristic attentional shift when problem-solvers reach the solution of Tangram, eye tracking data was recorded during the participants manipulated the 7 pieces. Areas of interests (AOIs) were surface of the each piece and the silhouette which was presented to the participants. When the participants looked at the range of a single piece or at the silhouette for more than 33 while manipulating the pieces, this duration time was accumulated as a time spent looking at the AOI. The proportion of the time spent looking at the AOI per second during the total manipulation time was calculated for the each piece, and it summed over the 7 pieces for each participant. Table 3 shows the mean proportions of the time spent looking at the pieces and that looking at the silhouette, for the high-achievers who completed both the tasks and for the non-achievers who could not complete the both tasks. Statistical analysis revealed that the proportion of the time spent of the non-achievers looking at the pieces (283.1 msc/sec) was significantly longer than that of the high-achievers (137.4 msc/sec), $F(1, 13) = 10.9$, $p < .01$. There was no difference between these two groups in the proportion of the time spent looking at the silhouette. The result that the time spent looking at the range of the pieces was relatively short might be reflected a wider or more active scanning over the problem space. In order to verify this inference more directly, eye movement distance during the participants manipulated the piece was calculated. The values in the bottom row of Table 3 indicate a mean eye movement distance per second in a visual angle for the high-achiever and for the non-achiever. Statistical test revealed that the eye movement distance of the high-achiever was higher than that of the non-achiever but it was marginally significant, $F(1, 13) = 3.15$, $p < .10$.

Table 3: Mean proportion of the time spent (msc/sec) and the eye movement distance (deg/sec)

| | High-achiever | Non-achiever |
|---|---|---|
| Proportion of time | | |
| 7 pieces | 137.4 (40.0) | 283.1 (108.3) |
| Silhouette | 25.7 (31.4) | 13.2 (11.9) |
| Movement distance | 72.4 (47.3) | 36.8 (25.3) |

## Discussion

This research aimed to demonstrate a constraint relaxation which is followed by the transition to an appropriate representation in insight problem solving. For this purposes, Tangram was used as a new tool. To investigate the correspondence between the constraint relaxation and actual manipulation of the pieces of Tangram, combinations of the 2 largest triangle pieces which were the key to completing the two task silhouettes, a lion and an arrow, were analyzed. Nakano (2009) found that problem-solvers of Tangram tended to combine these 2 pieces into a geometric shape. This initial constraint would facilitate reaching an insight for the solution of the arrow task, because the correct configuration of this silhouette could be achieved by arranging the 2 pieces into either a square or a triangle. As expected, in this task both the completers and the non-completers arranged the 2 pieces predominantly in a geometric combination through the problem solving. While in the lion task, this initial constraint to arrange the pieces into a geometric combination would inhibit the insight to occur, because the 2 pieces should be arranged in an irregular combination to complete the correct configuration of the lion task (top-right in Figure 2). Thus, the participants were expected to achieve better results in the arrow task than in the lion task. Contrary to this prediction, however, neither the percentage of the completers nor the time to completion of the high-achievers who were completed the both tasks were significantly different between the two tasks. This finding supported the view that the relaxation of the initial constraint was not the sole determinant of the insight for reaching the solution (Ormerod et al., 2002). As an evidence of this interpretation, in the lion task, the percentage of the cumulative duration time that the 2 pieces were arranged in the geometric combinations was decreased in the second half comparing with the first half. Importantly, such the decrease was found not only in the completers but also in the non-completers.

The critical difference between the completers and the non-completers was found in the lion task, in that, the completers arranged the 2 pieces in the irregular combination for a longer time in the second half than in the first half. This result for the completer should be a manifestation that the relaxation of the inappropriate initial constraint was followed by the construction of more appropriate representation. In contrast to this steady approach of the completers to reach the solution, the non-completers could not distinguish which of the three types of combinations would lead them to the solution, even in the second half. Therefore, the critical determinant for reaching an insight to the solution was a clear cut differentiation between the appropriate representation and other alternatives.

The participants' eye movements were measured while they manipulated the pieces, in order to investigate specific feature of attention shift that facilitated the successful problem solving in Tangram. The analysis of the time spent looking at the surface of the pieces indicated the noteworthy difference between the high-achiever and the non-achiever that the latter had a longer time staying attention on the pieces. In contrast to this attentional feature of the non-achiever, the high-achiever had a slightly longer distance of eye movements during the manipulation of the pieces than the non-achiever. These findings suggested that the successful problem-solvers of Tangram tended to search the problem space of the pieces more widely or actively.

The second purpose of this research was to demonstrate the independence between the transition from the initial but inappropriate constraint to the more appropriate representation for the solution and the changing of the confidence to reach the solution. There was a remarkable difference between the results of the present experiment and the series of the researches by Metcalfe (Metcalfe, 1986; Metcalfe and Wiebe, 1987). In these previous researches, most of the participants rated their confidence lowest level at the starting of the problem solving, and their self-evaluations stayed constant or slightly increased by the floor effect. While in the present experiment, the evaluation to the confidence was relatively high before starting the problem solving of Tangram, and it declined over time. Considering that the findings of Metcalfe was obtained by using insight problems including a spatial task and a linguistic task, Tangram is more likely to give the impression that it is easy to solve than those other insight problems, especially at the beginning of the problem solving. Another notable finding was that the participants who solved the insight problem had evaluated less confidence to reaching the solution (Metcalfe, 1986; Metcalfe and Wiebe, 1987). In other words, the actual achievement of each participant had opposite direction to their subjective confidence. Furthermore, the present experiment indicated that the completer manipulated the pieces so as to approach to the solution steadily but their confidence about it consistently declined. Therefore, this finding demonstrated that the process of deriving an insight and the subjective awareness to the problem solving did not progress independently but in the opposite direction over time.

## Acknowledgments

This research was supported by JSPS KAKENHI Grant-in-Aid for Challenging Exploratory Research (Grant Numbers 26540060).

## References

Jones, G. (2003). Testing two cognitive theories of insight. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 1017–1027.

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology*, 22, 374–419.

MacGregor, J. N., Ormerod, T. C., & Chronicle, E. P. (2001). Information processing and insight: A process model of performance on the nine-dot and related problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 176–201.

Metcalfe, J. (1986). Premonitions of Insight Predict Impending Error. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 623-634.

Metcalfe, J., & Wieb, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition*, 15, 238-246.

Nakano, Y. (2009). The insight problem solving in the puzzle game of “Tangram”. *Memoirs of Faculty of Education and Human Studies Akita University: Educational Sciences*, 64, 65–72.

Ormerod, T. C., MacGregor, J. N., & Chronicle, E. P. (2002). Dynamics and constraints in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 791–799.

Thomas, L. E., & Lleras, A. (2009). Covert shifts of attention function as an implicit aid to insight. *Cognition*, 111, 168–174.