3/28/2014

Variation, impasse and insight

Anna Fedor, Michael Öllinger, Eörs Szathmáry

Parmenides Foundation, Munich, Germany

Variation, impasse, and insight

Authors: Anna Fedor, Michael Öllinger, Eörs Szathmáry

Affiliation: Parmenides Foundation, Munich, Germany

1 INTRODUCTION

Most of us know how it feels, to get stuck when we puzzle over a difficult problem. There is a feeling, of no further progress, no options, and no idea how to continue. Despite of anecdotal evidence, it is still controversial whether an impasse is a necessary condition for having an insight. Accordingly, there is no consensus whether insight problem solving is something special (Davidson, 2003). The proponents following the Gestaltist (Duncker, 1945; Köhler, 1925; Wertheimer, 1959), that demarcated productive thinking from “blind” reproductive thinking. The basic process is restructuring that provides new insights in the given problem and allows a solution. Nowadays, researchers are interested in the underlying cognitive processes and in the neural correlates of insight problem solving (Kounios & Beeman, 2014; Luo & Niki, 2003) that might indicate that insight is actually different from standard problem solving (Metcalfe & Wiebe, 1987). It is still an unresolved question how people find an insight and what processes contribute.

The idea to understand problem solving by consecutive phases that are distinguishable by overt behavioral changes is an old one. As far back as the beginning of the 20th century, Wallas (1926) suggested four different stages in creative problem solving i) preparation, ii) incubation, iii) illumination, iv) verification. In the incubation phase is where active problem solving attempts halt and illumination means that an insight unexpectedly occurs. Finally, verification is necessary to prove that the putative insight actually solves the problem.

1.1 IMPASSE AND INSIGHT

More recently, Ohlsson (1992) stressed the importance of impasse for insight problem solving. According to him, an impasse is crucial for initiating a representational change that relaxes over-constrained problem representations. That is, Ohlsson defined impasse as a mental state that affects problem solvers’ solution behavior, and he proposed that insight can be defined as breaking out of an impasse (p. 4). The importance of this assumption is that it overcomes the prevailing circular definition of insight problems being problems that are solved by insight, or phenomenological accounts that pinpoint insight by having an Aha! experience, being sudden, discontinuous, unintended, and unexpected (Metcalfe & Wiebe, 1987; Öllinger & Knoblich, 2009; Sternberg & Davidson, 1995; Weisberg, 1995). Ohlsson outlined a process definition where insight is intimately connected with impasse.

Seifert et al. (1995) provided a multi-stage model to explain insight problem solving. Within the model impasse plays an important role for storing "failure indices" – i.e., a history of unsuccessful solution attempts is stored. They assume an impasse as crucial for guiding problem solvers towards further solution attempts. An intriguing question for them was, how people realize that they get stuck. They speculated that two criteria might be important for recognizing an impasse. First, people might come up with estimations on the required time a problem will need, consequently, if the allotted time elapsed people might believe they get stuck. Second, when people realize that all available components, and strategies of the problem are exhausted, they feel stuck.

Beeftink and colleagues (2008) differentiated between impasse and fixation (Öllinger, Jones, & Knoblich, 2008). Impasse was characterized by the feeling of being stuck and by affective states like frustration and confusion. The authors assumed that fixation precedes impasse. Fixation does not immediately lead to the realization that a problem is difficult and have no obvious solution, but let participants continue with solution attempts that fail to solve the problem and that become more and more frustrating until they get stuck.

Öllinger and colleagues (Öllinger, Jones, Faber, & Knoblich, 2013; Öllinger, Jones, & Knoblich, 2013a, 2013b) elaborate on Ohlsson's model and found e.g. for the nine-dot problem (Öllinger, 2013a) that before an impasse, successful solvers used heuristics that restrict the initial over-constrained search space. These help them to realize that the initial search space is inappropriate (because it is over-constrained) to find the solution – so they get stuck in an impasse. Within an impasse a representational change occurs by releasing the constraint: participants start to draw lines outside the emerging square-shape of the nine-dots, and again, heuristics are crucial to restrict the now even larger search space to come up with the solution. Complementary, MacGregor and colleagues (2001) proposed a computational model for the nine-dot problem suggesting that impasse is reached, when no further progress can be made by the selected criterion for progress (e.g. connecting as many dots as possible with each available line), and impasse is broken, when such a criterion failure initiate search for alternative moves.

There are a few attempts to find explicit evidence for the existence of impasses in insight problem solving. Knoblich et al. (2001) used eye-movement data to identify different problem solving phases. They used matchstick arithmetic tasks that require turning an incorrect arithmetic statement into a correct one, by moving one single matchstick (Chi & Snyder, 2011; Knoblich, Ohlsson, Haider, & Rhenius, 1999; Öllinger et al., 2008; Reverberi, Toraldo, D'Agostini, & Skrap, 2005). Simple matchstick arithmetic tasks require changing the values of the incorrect equation. However more difficult and insightful solutions require the problem solver to change the operators of the equation. For the latter Knoblich and colleagues (2001) identified three different eye-movement patterns that they assigned to different problem solving phases. Initially, all participants preferred to look significantly longer at values of the equations. In a second phase, unsystematic eye-movements were found, according to the authors an indicator for being stuck in an impasse. After overcoming an impasse, successful solvers changed their preference to gaze significantly longer on the operators of the equations (Knoblich, Öllinger, & Spivey, 2005). Jones (2003) recorded eye-movement data during participants solving versions of the car park problem and he identified two phases: before an impasse, and during an impasse. Before an impasse he identified the importance of heuristic search processes (like hill-climbing), whereas within an impasse representational changes can take place.

Sandkühler and Bhattacharya (2008) proposed clearly distinguishable neural correlates for different stages during insight problem solving (see Kounios & Beeman, 2014 for a critique of this approach). The authors found that different EEG frequencies could be assigned to i) mental impasse, ii) restructuring, iii) deeper understanding of the problem, iv) AHA! feeling that stands for the suddenness and obviousness of a solution.

Beside these explicit impasse accounts, there are approaches that implicitly differentiate between phases. E.g. in their seminal work Metcalfe and Wiebe (1987) asked participants in regular intervals to rate how far they believe they are from the solution. They called this meta-cognitive measure "feeling of warmth rating" and found that for insight problems people were unable to estimate the distance, except when they are right before the solution. On the other hand, participants performed very well when rating non-insight problems that showed a linear increase of warmth.

On the contrary, Fleck & Weisberg (2013) and Kounios & Beeman (2014) accept the importance of representational change, but they deny the necessity and the existence of the impasse-insight sequence. Fleck and Weisberg (2013) found using detailed analyses of thinking aloud protocols that in general insight problems can be solved by various methods and that the "'classic' impasse-restructuring-insight sequence occurred in only a small minority of solutions" (p. 436). Consequently, they argued that there is no clear or "sharp" criteria that differentiate between insight and non-insight solutions, least of all impasse.

Recently, Kounios and Beeman (2014) argued that it is not justified to use impasse as defining feature for insight, because it would exclude various other types of insights, like when a solution pops into someone's mind although she is engaged in analytical problem solving without already reaching an impasse.

Taken the ambivalent evidence, it seems indeed plausible that some people show the proposed "before impasse-impasse-restructuring-after impasse sequence", but also that it is not a necessary condition for having an insight, and most important no clear cut-off criterion that differentiate between insight and non-insight problem solving. A question that can be asked is whether the predominant definition of impasse according to Ohlsson is the proper one. We aim at elaborating on this question and attempt to clarify the process of insight problem solving and the notion of insight. To reach this goal we suggest new classifiers on people's problem solving data that were lent from the evolutionary Darwinian framework.

A question that directly follows is whether constantly "doing something" (Fleck & Weisberg, 2013; Kounios & Beeman, 2014) is actually an indicator for being not in an impasse. We suggest that it might be worthwhile to introduce a clearer criterion that differentiates between potential problem solving stages.

1.2 THE PRESENT STUDY

For the current study, we used a version of the Katona problem (Katona, 1940; Öllinger, Jones, & Knoblich, 2013a, see Figure 1). The task is to reduce the number of the given five squares to four squares of equal size by moving only three sticks, without discarding any stick. The problem proved fairly difficult and Öllinger et al. (2013a) showed that the main source of difficulty was caused by the necessity to decompose the given closed figure of five squares into two separate compartments. That is, the representational change is to relax the over-constrained assumption that a solution has to be a single entity.

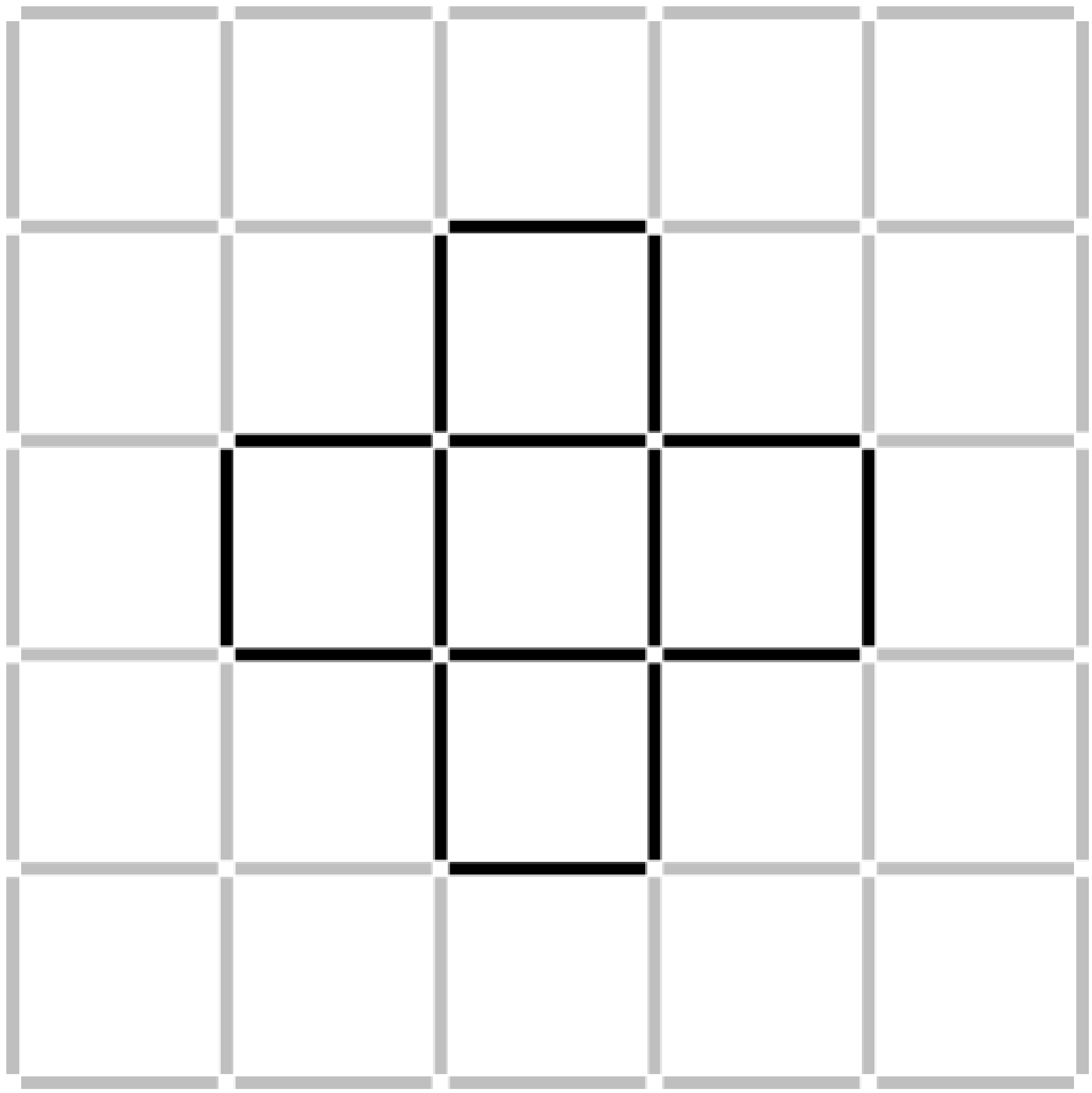


Figure 1. The initial position of sticks on the grid in the Katona problem. For the solution, see the Appendix. We show a 5-by-5 grid here, but in the computerized task, participants saw a 9-by-9 grid.

1.3 Predictions

We are interested how the selected matchsticks and the corresponding end positions vary across the solution attempts. We will see that initially, participants attempt to manipulate different sticks and visit different end positions. After a while the change rate stops increasing and reaches a plateau – this means that the participant makes moves that uses end positions that were already used previously. For those, who solve the task, this plateau usually ends right before they find the solution, which, according to the process-based view of insight means that the plateau could be the indicator for being within an impasse. The implication is that although participants continue with overt and tentative move behavior they make no further progress (Chronicle, MacGregor, & Ormerod, 2004; MacGregor et al., 2001; Ormerod, MacGregor, & Chronicle, 2002). We would suggest classifying this kind of behavior as an indicator of impasse. We predicted a high correlation between the plateau and the subjective feeling to be stuck in an impasse. This prediction was not strengthen by the data. The beauty of analyzing the moves of participants in this mechanistic, bottom up way is that it can pinpoint problem solving stages objectively.

2 METHODS

2.1 THE TASK AND DESIGN

We used one version of the Katona problem (Katona, 1940), the cross shape (see Figure 1), in which the task is to make four squares of equal size by moving exactly three sticks. We introduced two experimental groups. In the impasse monitoring group (IMG) participants were provided with an instruction of the impasse feeling, and they were asked to press a button if they felt they got stuck. A second group served as a control group (CG) to control whether the metacognitive task of monitoring one's impasse state had an impact on the problem solving process.

After the task was finished, either solved or not, all participants (Ps) were asked whether they felt insight.

2.2 MATERIALS

2.2.1 Data collection

We provided a computerized version of the Katona problem, where participants (Ps) could move single sticks by drag and drop with the computer mouse. The program was written with Microsoft Visual Basic ® and run individually on the participants own computers. We provided a downloadable version of the program on the internet. The program recorded the stick positions and the movement of the sticks and the button presses along with the time that passed since starting the task. It has automatically checked whether a correct solution was achieved. The data was stored in a *.csv file.

2.2.2 Procedure

Ps were tested individually. The program guided Ps autonomously through the experiment. To familiarize Ps with the drag and drop mouse movement, the program started with a practice trial, where participants had to drag and drop four sticks on the screen to four different grid positions. After completing this task Ps received the following written instruction:

> *You will see five squares made of sticks on the screen. Your task is to move exactly three sticks to produce four squares of equal size, while leaving no sticks that are not part of squares.*
>
> *After three moves the task resets automatically or you can reset it anytime by pressing the Reset button. You can try to solve the task as many times as you like, in fact, try to show us all your ideas by moving the sticks, don't just try to solve the task in your head. You will have 15 minutes to solve the task – try to use all of it. It is not a problem, if you cannot solve the task, but if you close the program before 15 minutes elapsed, your submission will be invalid.*

Ps of the self-monitoring condition received the additional instruction:

> *Sometimes before solving a difficult problem people feel like they are stuck, they are not getting closer to the solution. We would like to know if you feel like this during*

the task, so in this case please press the "I'm stuck" button. You can press the button more than once if you feel that this feeling increases.

Ps were allowed to reset the initial configuration of sticks as often as they wanted. After an unsuccessful attempt the program provides feedback and resets automatically. After solving the problem or reaching the upper time limit of 15 minutes, participants were asked to copy the output file and post it to the experimenter via our website and complete a short online questionnaire (age, gender, handedness, vision, educational background, nationality, mother tongue, level of English). They were also asked whether they solved the problem with or without insight:

Some people feel a sudden, unexpected, unintended, and surprising moment where a solution pops into someone's mind. The accompanying experience is often called "Aha!" experience. Did you have this feeling before or when you solved the task?

The P had to choose between the following answers: I did not solve the task; No, I did not feel anything like this; Yes, I felt exactly like this; Other: [free text].

2.3 PARTICIPANTS

2.3.1 Recruitment

Most of the Ps were recruited via the internet, on a crowdsourcing platform called CrowdFlower (four Ps were recruited at a university). Here, people willing to work online can sign up and then choose from the available jobs that employers offer. These crowdsourcing platforms are usually used by employers to get those kinds of jobs done that are easy to serialize, but are not easy to automatize, like CAPTCHA recognition, picture sorting according to content, etc., or to do market research via questionnaires. They are also more and more often used to do online experiments.

We recruited level 3 workers, who are the highest rated, most trustworthy group of workers on CrowdFlower. They received one dollar as a payment. The participants were randomly assigned to either the impasse monitoring group or the control group.

2.3.2 Sample size

We had more than twice as many Ps in the IMG than in CG because we wanted to analyze IMG in more detail, namely perform a paired t test on those Ps in IMG who pressed the impasse button. We computed the required sample size for this test *a priori* with the G power software, assuming a medium effect size (0.5) and setting $\alpha = 0.05$ and power = 0.8. According to the analysis, the sample size for a matched-pairs t test with these parameters should be 34, which we multiplied by 2.5 thinking that about half of the Ps won't be able to solve the task and not all of those will press the impasse button. At the end, this analysis proved to be insufficient, since we assumed normal distribution, but the data was not normally distributed, so we had to do a nonparametric test instead of the planned paired t test.

2.3.3 Exclusion criteria

We excluded all participants who closed the software before either they solved the task or the 15 minutes elapsed, even if they restarted the software afterwards. We did this, because we did not know how much time elapsed between they closed it and restarted it and we did not know whether they spent this time trying to figure out the solution off-line.

2.3.4 Demographics

After excluding Ps who closed the software early and/or did not finish the task, we have left with 129 Ps; 42 in Group 1 and 87 in Group 2. There were 27 females and 102 males. The average age of Ps was 30.3 years (range 16 - 69). We had Ps from 37 countries, five continents (Europe = 59, Asia = 44, North-America = 9, South-America = 12, Africa = 5). The solvers and non-solvers were almost equally distributed across continents. Only four countries had more than five Ps (India: 20, USA: 9, Romania: 7, UK: 6). Most of our Ps had some higher education (**Errore. L'origine riferimento non è stata trovata.**). 11 Ps had basic levels of English, the rest judged their English as intermediate level or higher.

3 RESULTS

3.1 DATA ANALYSIS, DEFINITIONS

First, we have to define a few concepts that are necessary for our analyses. The grid defines the search space of potential vertical and horizontal stick positions. A *position* is a location on the grid, where sticks could reside. It could be a position where the sticks were originally placed, or a position that was empty at the start of the task. A *move* involves picking up a stick and putting it down somewhere else. If the P picked up a stick and then put it back into the same position, it was not considered a move. The position from where the stick was picked up is called a *source position*, a position where the stick was put down is called a *target position*. A target position could serve as a source position in a later move and a stick could be put back into a position where there were sticks at the start of the task.

A *trial* consists of maximum three consecutive moves. After three moves a new trial begun automatically. A trial could consist of less than three moves if the P hit the “Restart” button before completing all three moves. A *candidate solution* is a list of source and target positions in a trial. It is composed of two, four or six positions, not taking into consideration the order of moves and the link between the source and target positions. It is usually sufficient for reconstructing the final configuration of sticks in a trial.

As can be seen in Figure 1, the initial configuration of sticks is symmetric. Consequently, there are four different sites where a solution could occur and the symmetric moves are equivalent. For our analyses, we decided not to take into consideration the symmetry of the figure and when we count the number of different moves a P makes, we do not collapse equivalent moves into the same category. We did this, because it turned out that Ps did not use all segments of the figure equivalently: the top left portion was used less often than the rest.

We used an alpha level of .05 for all statistical tests.

3.2 EFFECT OF MANIPULATION

3.2.1 Solution rate

To control whether the additional metacognitive task in IMG affected the problem solving performance of Ps, we compared the CG and the IMG in terms of solution rate and solution time.

The overall solution rate was 51%; 54.8% in CG and 49.4% in IMG (Table 1). According to a Chi-square test, the solution rate was not influenced by the additional task of having to monitor one’s feelings and occasionally pressing the impasse button: X^2 (1, N = 129) = .14, p = .70.

Table 1.

Solution rate of participants			
Participants	**CG**	**IMG**	**Total**
Solvers	23	43	66
Non-solvers	19	44	63
Total	42	87	129

3.2.2 Solution time

29 Ps solved the task under 1 minute in the two groups. The rest of the solution times were more or less evenly distributed between 2 and 15 minutes. For non-solvers, the task always ended after 15 minutes.

Looking at the solution time of solvers in the two groups separately (Figure 2), although in the CG it is spread higher, but the medians are close, and the difference between the groups is not significant (Mann-Whitney U = 424, p = .3462).

Since we did not find differences between the groups, we merged them for some of the following analyses.

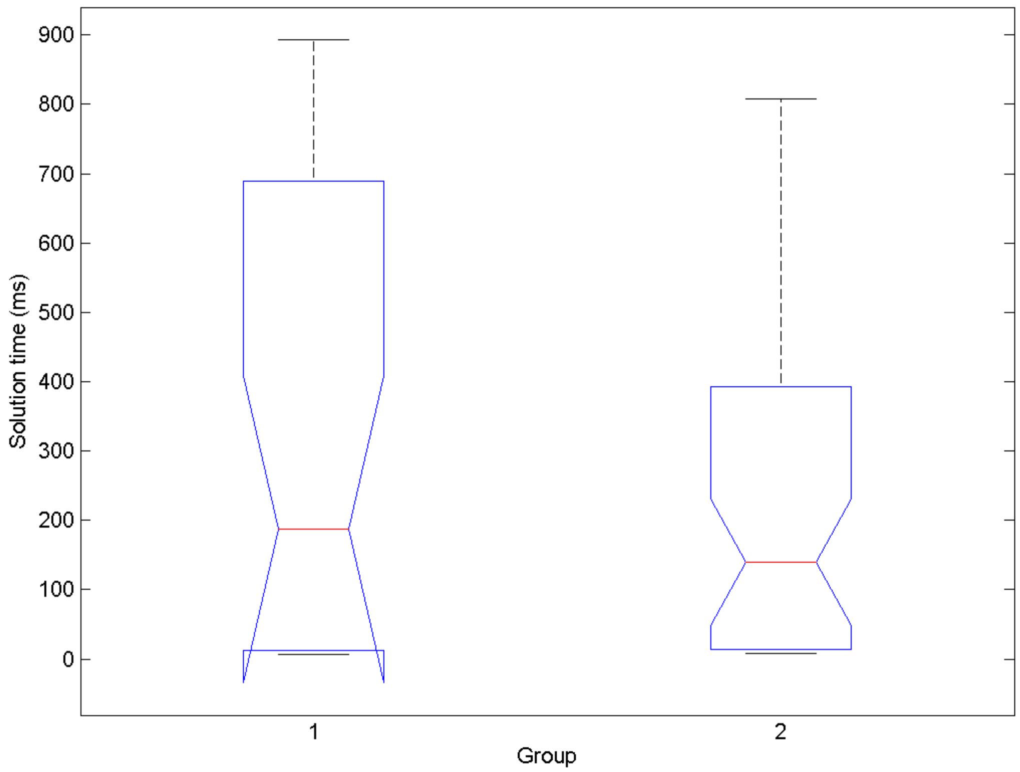


Figure 2. Comparison of solution times in CG (1) and IMG (2). The central mark is the median, the edges of the boxes are the 25 and 75 percentiles and the whiskers extend to the most extreme data points not considered outliers. Outliers would be data points > $q_3 + 1.5(q_3 - q_1)$ or < $q_1 - 1.5(q_3 - q_1)$. The extremes of the notches show 95% confidence interval of the median.

There were 28 Ps who only had only one trial. 26 of these solved the task under 32 seconds; one solved it in 3.4 minutes and another one did not solve it. It is probable that the former solved the task in his head and the latter gave up after only one trial, and those 26, who solved it in about half minute either already knew the task or just found the solution instantly.

3.3 Number of grid positions used

We looked at different measures of variation of solution attempts. The most basic of these is the number of different grid positions that were used by the P, either as source or as target positions. We counted the number of positions after each move in a way that each position was only counted once; thus a move could add 0, 1, or 2 new positions to the pool of already used positions.

3.3.1 Plateau

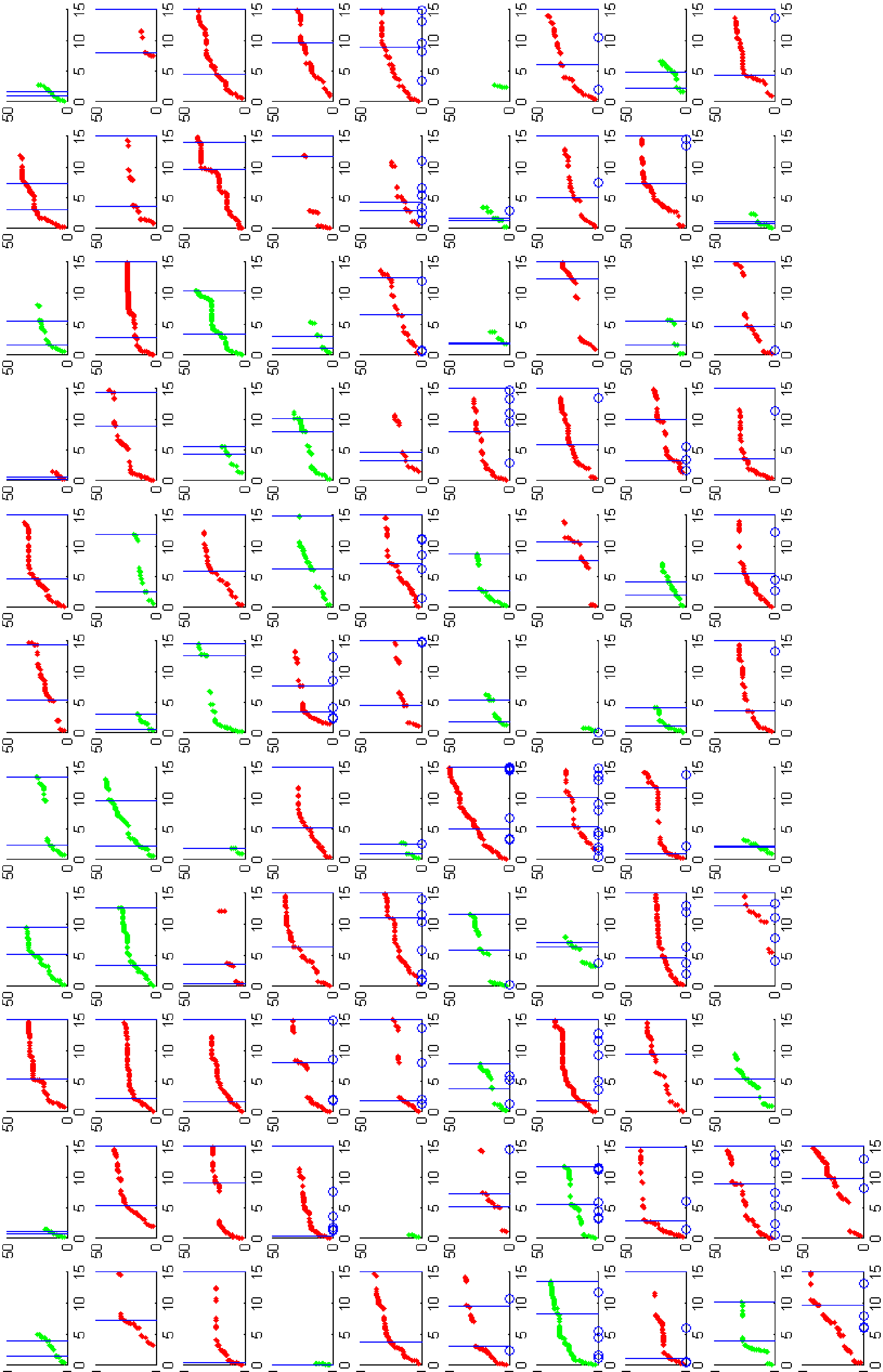
We plotted the number of used positions against time (Figure 3). The number of used positions measures the search space that the P considers for the solution, or at least a sample of it that manifests in moves. If we look at the plots P-by-P, a general pattern emerges. For many Ps, at the beginning of problem solving, the number of positions increases with every move. After a while it stops increasing and reaches a plateau – this means that the participant makes moves that uses grid positions that were already used previously, i.e., she repeats herself. This plateau usually begins few minutes into the task and ends either at the end of the problem solving (i.e., at timeout, or when the P solves the task), or a little bit before that. Of course, not all plots conform to this pattern: sometimes there is no plateau, and sometimes it starts at the very beginning of problem solving. But in general, the plateau divides the curve into two or three sections: before plateau, plateau, and after plateau, and if one of these are missing, it is usually the after plateau phase.

To further investigate this pattern, we defined the plateau in the following way: the *plateau* is the longest stretch of moves where none of the moves increases the number of positions by more than 1. A move, in which both the source and the target positions are new, compared to the previous moves, breaks the plateau. If there are more than one plateaus, we choose the longest.

We would like to argue that the plateau in the size of the search space indicates the impasse stage of problem solving. To corroborate this hypothesis, we wanted to see whether common assumptions about the impasse phase are also true for the plateau. Of course, there are no assumptions that are true for every problem solver, but in general, we assume that:

- The impasse phase is preceded by a deliberate search phase
- In the case of solvers, the impasse phase is followed by restructuring, causing an insight that entails a probably short and strongly directed search phase
- Non-solvers are stuck at the impasse stage, i.e., it lasts until the end of problem solving (15 minutes)
- At the impasse stage, problem solvers feel that they are stuck

Figure 3. The number of different grid positions used by participants. Each plot belongs to a participant in CG or IMG. We excluded those participants who solved the task in three moves (in one trial). The horizontal axis shows time in minutes, the vertical axis shows the number of grid positions. Each data point represents a move. Solvers are plotted by green, non-solvers are plotted by red. Blue horizontal lines indicate the start and the end of the plateau identified as described in the text. Blue circles on the horizontal axis indicate time points when the participant pressed the impasse button in IMG.



3.3.2 Problem solving stages

To test the first three of these assumption, we counted how many problem solving stages emerge when we define the plateau, i.e., is there a phase before and after the plateau. Theoretically, the plateau could start at the second move, and last until the last move of the P. Alternatively, the curve could have a section with higher slope before and/or after the plateau. Since, by definition, there is only one plateau, there are no other possibilities.

Those 28 Ps, who solved the task in just three moves are not relevant for this analysis, so they were excluded. Apart from them, there were 3 more Ps, who solved the task and had no plateau in the size of their search space. We could identify a plateau in all of the non-solvers' plots. The plateau started at the first move in one P, i.e., there was no stage before the plateau. Two additional P had just one move before the plateau, but for the rest of the Ps, the first stage was longer than this.

The most interesting question is, whether the third stage tends to be missing in non-solvers rather than in solvers. The third stage is missing in 15 out of 39 solvers (38.5%) and in 46 out of 62 non-solvers (74.2%). A Chi-test revealed a significant difference, $X^2 = (1, N=101) = 11.33$, $p = .0008$.

On average, the first and second stages consist of more moves in non-solvers than in solvers and the third stage is shorter, in fact it has a median of 0 (Figure 4).

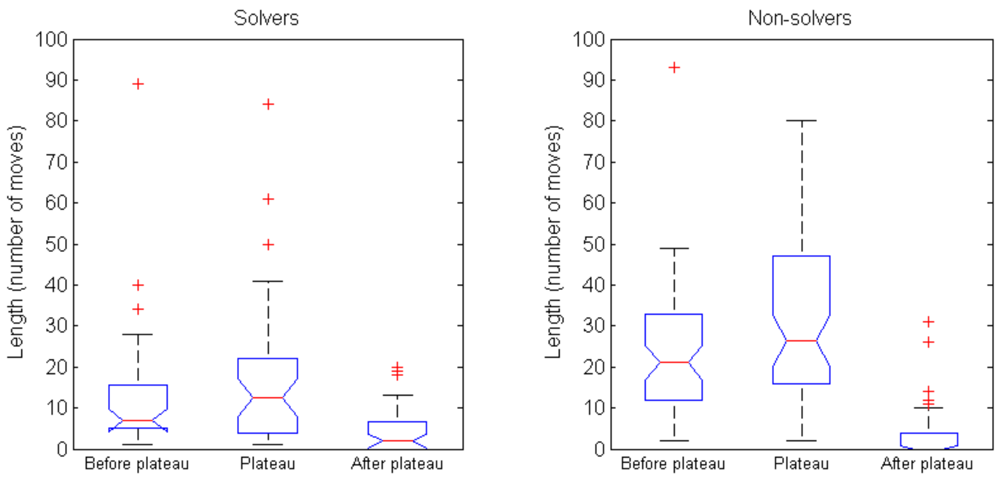


Figure 4. The median number of moves in each problem solving stage in solvers and non-solvers. Two outliers above 100 were cut off from the plateau stage of non-solvers.

We also looked at the length of problem solving stages in minutes. It shows a similar pattern as the length of problem solving stages in moves (Figure 5).

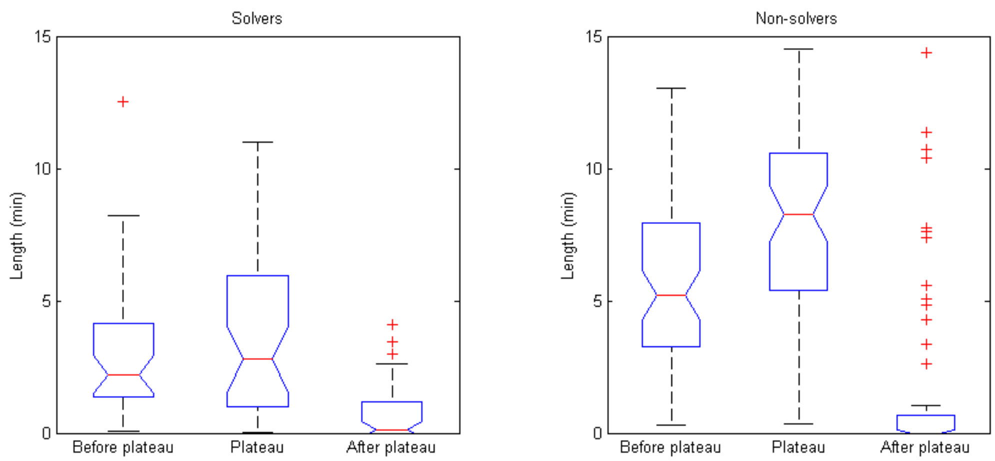


Figure 5. The median length of problem solving stages in minutes.

3.3.3 Plateau and impasse feeling

We wanted to see whether the plateau, which we believe to overlap with the impasse stage, also coincides with subjective impasse ratings. We counted the number of times the impasse button was pressed by Ps in each stage. 8 of the solvers and 35 of the non-solvers pressed the impasse button at least once. For them, we added the number of button presses outside the plateau and the number of button presses during the plateau and normalized it by the length of the corresponding stages (Figure 6). According to the Wilcoxon matched pairs signed-ranks test the difference between the median of the normalized number of button presses outside and during the plateau was not significant: $W = 145.00$, $p = .37$.

We looked at the average number of impasse button presses per minute separately for solvers and non-solvers (Figure 7). If there was an association between the supposed impasse stage and the subjective impasse feeling of Ps, we would see an increase in these curves towards the end of the problem solving time. We see no such increase, the curves oscillate randomly.

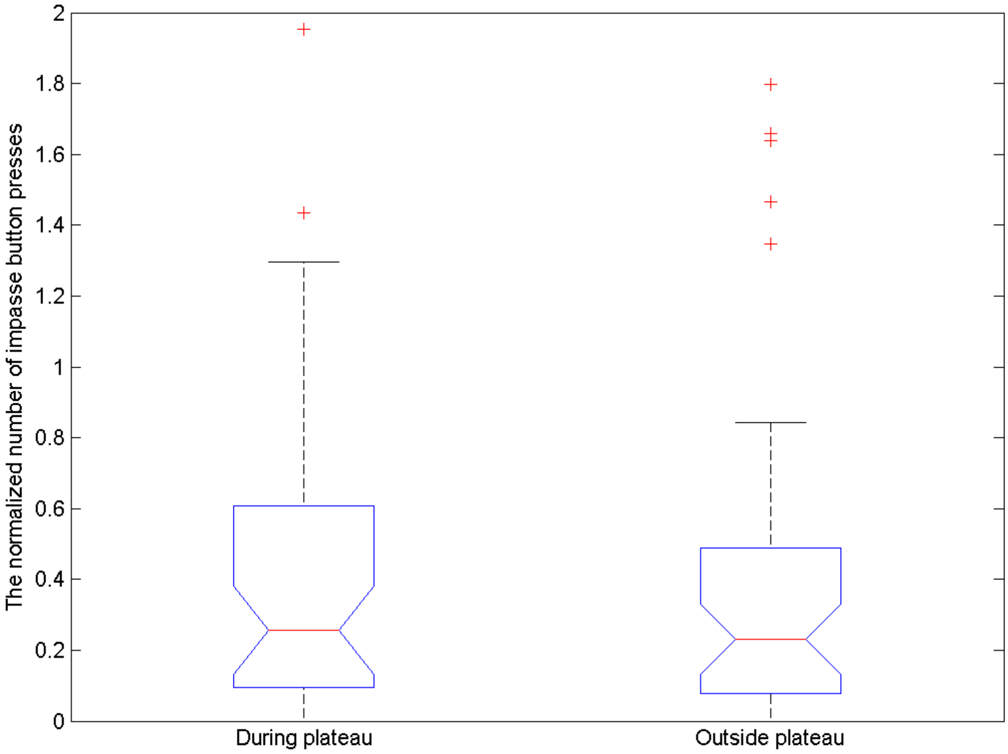


Figure 6. The median of the number of times participants pressed the impasse button during and outside the plateau, normalized by the length (in minutes) of the corresponding stages.

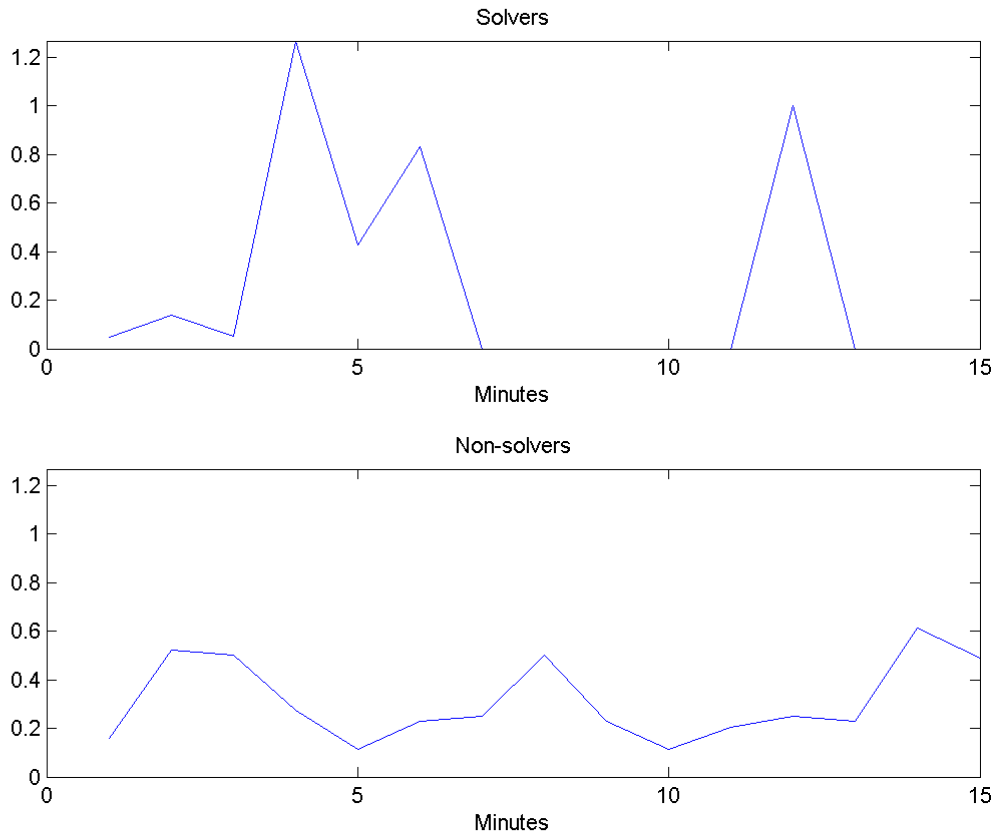


Figure 7. Average number of impasse button presses per minute.

3.4 NUMBER OF DIFFERENT MOVES

If we plot the number of different moves through time for each P, we see that the curves increase continuously – there is no plateau. This means, that although they use the same grid positions, Ps combine them into new moves.

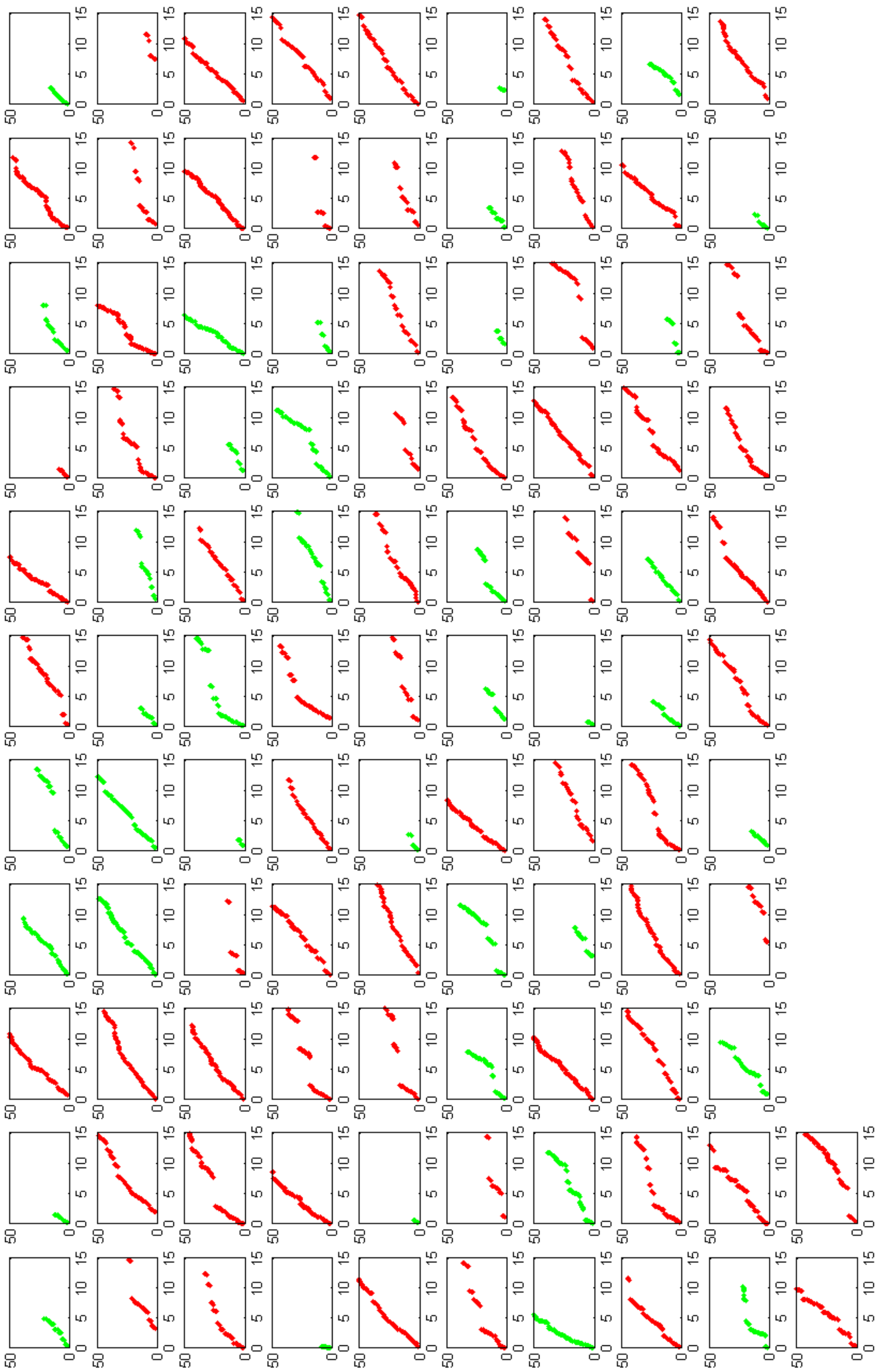


Figure 8.

participants who solved the task in three moves (in one trial). The horizontal axis shows time in minutes, the vertical axis shows the number of different moves. Each data point represents a move. Solvers are plotted by green, non-solvers are plotted by red.

3.5 NUMBER OF MOVES PER MINUTE

Some evidence (Knoblich et al., 2001; Ohlsson, 1992) suggests that people become inactive during the impasse. They feel like there is no solution, there are no good moves, it might even be impossible to solve the task, so they give up and problem solving attempts cease. In fact, looking at Figure 3, we can see gaps in the curve on many plots – these are stretches of time when the P did not make a single move. If the plateau in fact coincides with the impasse, and our general assumptions about impasse are true, we should see more of these gaps and less activity during the plateau than before or after the plateau (note, that the plateau was defined in terms of moves, not time, so this argument is not circular).

We compared the number of moves per minute before the plateau and during the plateau for those Ps who had a plateau (98 Ps) with Wilcoxon matched-pairs signed-ranks test and the difference was not significant: $W = 91.00$, $p = .87$. In spite of the obvious gaps in activity for some participants, the data on the group level does not support the hypothesis that Ps become less active during the plateau. On an individual basis, about half of the participants (53 Ps) were less active during the plateau, whereas the rest (45 Ps) were more active during the first phase of problem solving. We did not analyze the after plateau stage similarly, because it was missing in many Ps.

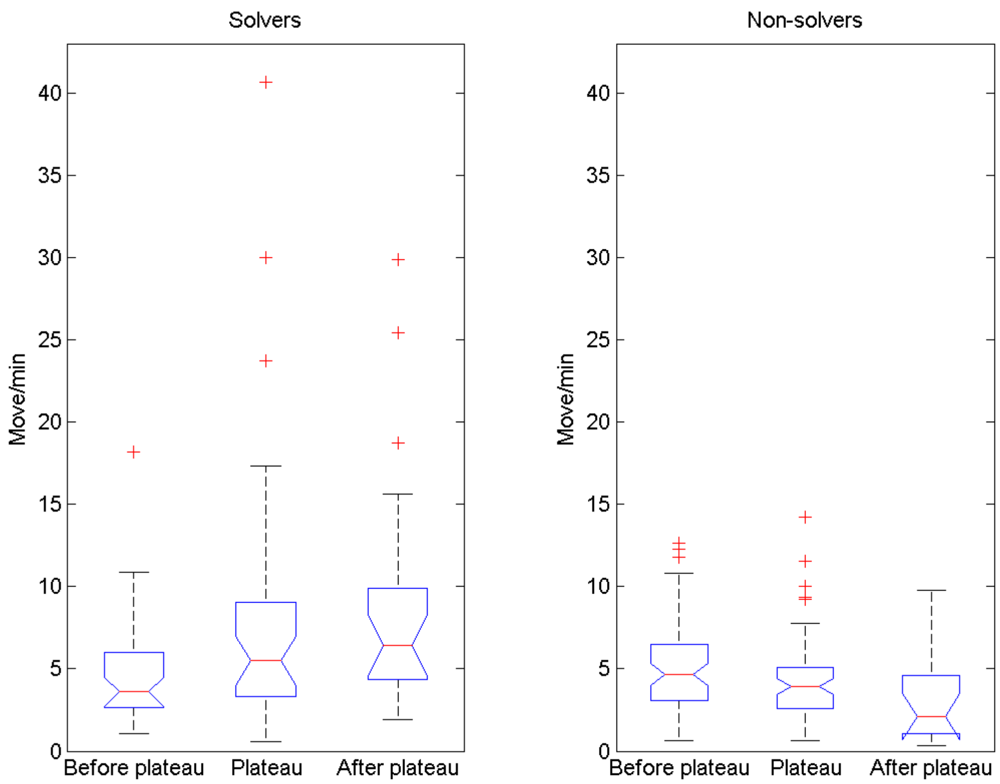


Figure 9. Average number of moves per minute for solvers and non-solvers in different stages of problem solving.

3.6 Solvers vs. non-solvers

3.6.1 Used positions on the grid

All 16 sticks were moved at least once by a P in the control condition. The most popular source positions were the sides of the inner square.

127 target positions out of the 180 possible grid positions were used at least once. Ps tended to use the middle of the grid, positions on the edges were rarely used. Most often, only less than a 5-by-5 grid was used in the middle of the screen. The most frequent target positions were the ones that complemented the cross to a big square, then the ones that complemented the cross to a cross with longer arms and some original stick positions. All in all, these positions form squares that have at least one side common with the original figure. In general, the top left corner of the figure was used less frequently than other portions.

3.6.2 Number of grid positions

We compared solvers and non-solvers on a group level, in terms of the number of grid positions they used (Figure 10). It seems that the curves highly overlap, at least until sample size becomes very small which in turn causes fluctuations.

Even though, on an individual level, we have seen differences between solvers and non-solvers (missing third phase for non-solvers, shorter plateau for solvers) on Figure 3, these differences disappear on a group level analysis.

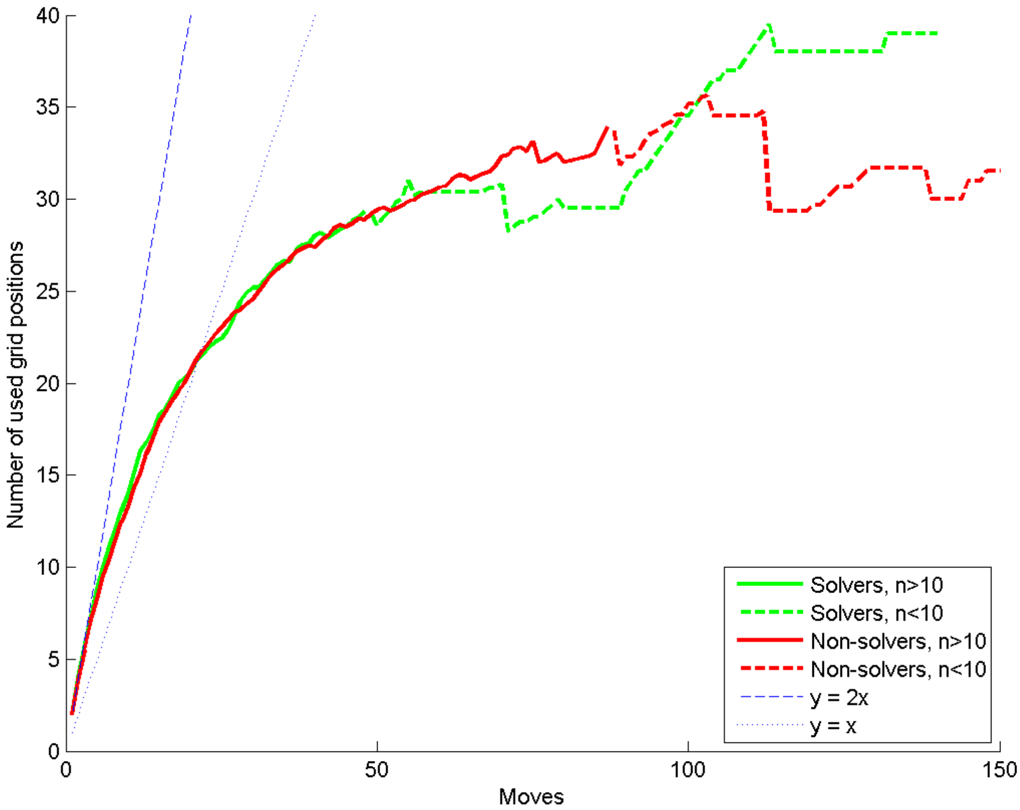


Figure 10. The average number of grid positions after each move.

3.6.3 Subjective feelings of insight and impasse

Of the 129 Ps in the two groups, six non-solvers (9.5%) and 46 solvers (69.7%) reported insight feeling after the task. Of those 26 solvers who solved the task under 32 seconds and in one trial, only 11 (42.3%) reported feeling insight (and none of them pressed the impasse button).

In the IMG, 43 of the 87 Ps reported feeling impasse at least once; of these 8 were solvers and 35 non-solvers (Table 2). Those, who pressed the impasse button usually did so less than 14 times (there was one P who pressed it 20 times and another who pressed it 59 times).

Table 2.

Number of participants who reported feeling insight and impasse				
Participants	**Insight**	**No insight**	**No impasse**	**Impasse**
Solvers	46	20	35	8
Non-solvers	6	57	9	35

3.6.4 Number of repeated candidate solutions

We have looked at how many times solvers and non-solvers repeated their candidate solutions. A repeated candidate solution contains exactly the same source and target positions as a previous candidate solution, but not necessarily in the same order (i.e., they might have been the results of different moves).

24 non-solvers and 11 solvers repeated a candidate solution at least once. The median of the number of repeated candidate solutions was significantly higher for non-solvers than for solvers (Mann-Whitney $U = 1623.5$, $p = .0273$). Also, non-solvers spent more time with the task and tried proportionately more candidate solutions. This could explain their higher repetition rate only partially; if we look at the number of repeated candidate solutions vs. the number of trials, the line fitted for non-solvers seems to increase faster than the line fitted for solvers (Figure 11).

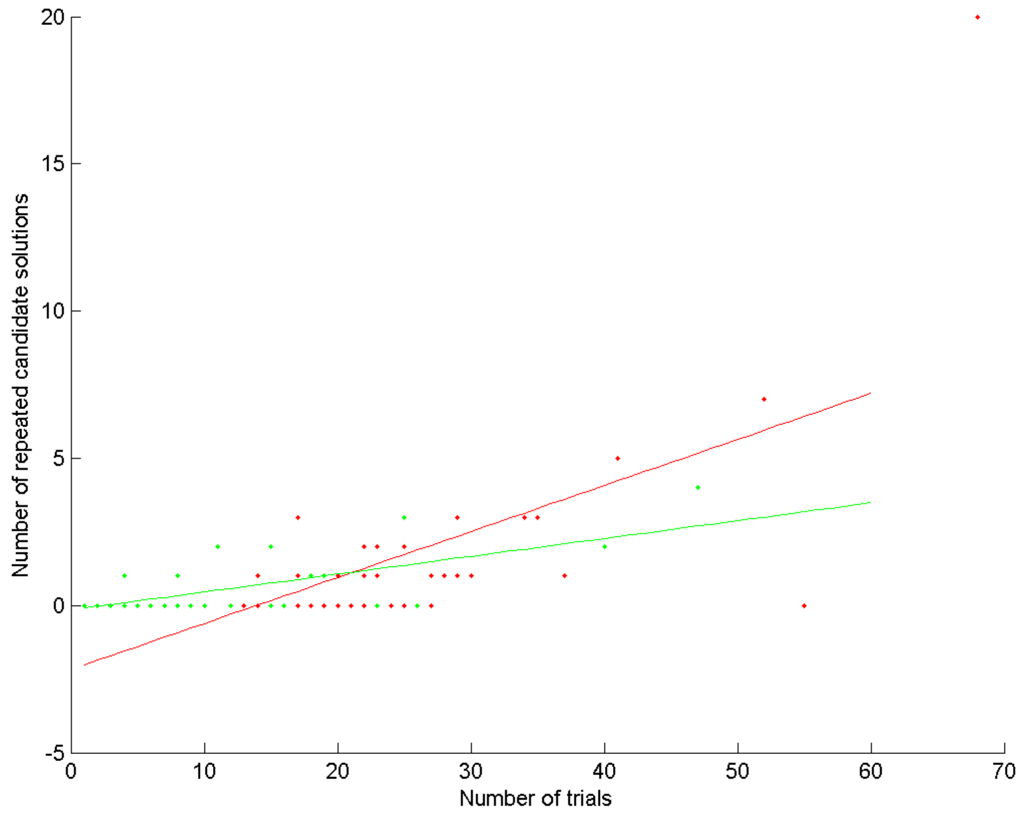


Figure 11. Number of repeated candidate solutions vs. the number of trials for each participant. Solvers are depicted with green dots, non-solvers are depicted with red dots. The lines represent linear regression was fitted separately for solvers and non-solvers.

4 Discussion

The current study aimed at clarifying the role of different problem solving stages during insight problem solving (see Öllinger et al. 2013a). At present, it is still unclear whether such stages exist and what are the fundamental characteristics of these.

4.1 Plateau and variation

We proposed that impasse can be elegantly operationalized on an individual level by analyzing move selection behavior. We introduced the concept of a plateau that indicates the stage when participants essentially do not generate new candidate solutions. This rigorous definition overcomes unreliable subjective self-reports, and complements coarser analysis of changes in the overt behavior. At the same time it adds the counter intuitive perspective that although problem solvers show continuous movement pattern they are still stuck, since they do not investigate new positions in the search space. This might shed light on the reasons why some people were able to solve the problem and others were not. It seems that overcoming the fixation on a dominant set of procedures (Lovett & Anderson, 1996; Öllinger et al., 2008) might be crucial for solving the problem. That is, insight problem solving has something to do with variation, selection, and retention of potential candidate solutions (Campbell, 1960; Simonton, 1995) that might bridge the gap between cognitive psychology and Darwinian explanations.

In this vein, our data demonstrated that non-solvers repeat their candidate solutions more than solvers on average. We can only speculate about the possible reasons of this difference. One of these reasons could be that solvers simply have higher working memory capacity, thus their build up a history of failures (Seifert et al., 1995) that informs and directs new attempts and make it less likely to repeat unsuccessful candidate solutions (a similar argument as MacGregor et al. 2001 suggested for the successful use of heuristics on insight problems, and the individual look ahead value) – a potential relationship that can be easily investigated by further studies.

Another likely possibility arises if we suppose that the candidate solutions that we can observe as a result of participants' behavior are samples from a larger population of candidate solutions that participants consciously or unconsciously entertain. The size of this underlying population could be smaller for non-solvers or the size could be similar in all participants, but the diversity is higher in solvers. This means that the number of different candidate solutions is higher in solvers, whereas non-solvers have more identical copies of the same candidate solutions. Of course, it is also possible, that there is no essential difference between solvers and non-solvers: solvers are just the lucky ones, who find the solution by chance, whereas non-solvers keep trying for longer, and it means that they inevitably bump into the same candidate solutions after a while.

4.2 Plateau and impasse

Unexpectedly, we found no relationship between the subjectively experienced impasse and the analytically determined plateau. It seems that participants are not aware of their restricted search space during the plateau. A possible reason for this could be that even though they repeat the source and target positions, they combine them into new moves. The result is the same, but the new combination of positions could mask the repetition of positions. The lack of awareness of the repetitions could account for the lack of impasse feeling.

Moreover, the subjective impasse reports do not aggregate at any particular stage during problem solving – they are evenly distributed through time. This means, that either the impasse stage does not exist, or it cannot be identified by subjective impasse reports. We believe that the latter is true and that the impasse stage can be identified by the plateau, when participants repeat their source and target positions. The plateau shows some of the properties of the previously rather elusive impasse stage, namely, it is usually preceded by a deliberate search phase, and solvers tend to get out of it at the end of problem solving, whereas non-solvers stuck in it until timeout.

We also found, when analyzing the distribution of reported impasses, that there was a significant difference between solvers and non-solvers. Significantly more of the non-solvers feel stuck than solvers, although the latter one showed plateaus, too. This finding is interesting in various aspects. First, it showed that the classical impasse definition (Ohlsson, 1992), feeling stuck, having no idea or clue, holds true particularly for non-solvers and reflect a realistic state of mind. On the contrary solvers might although get stuck by producing similar solutions (plateau), but have the feeling that they still continuing, with their solution attempts.

4.3 Limitations and further directions

Our approach is not applicable to all insight problems. The more moves participants have to make the better for our analysis. The algorithms we use need large amounts of data across time to produce meaningful results, so we can detect the existence of a plateau. That is, a large class of single move problems are not analyzable with our approach. Moreover, the analyses are naturally only applicable on overt solution attempts. Mental attempts and operations remain opaque in our model. On the other hand, our approach opens the door to investigate other classes of problems like implicit learning tasks or dynamic system tasks (Cleeremans, Destrebecqz, & Boyer, 1998), that usually are not considered as insight task, but as Haider & Rose (2007) suggested, provide a wide field, where the basic processes of rule detection, restructuring, and insight could be investigated.

5 References

Beeftink, F., van Eerde, W., & Rutte, C. G. (2008). The effect of interruptions and breaks on insight and impasses: Do you need a break right now? *Creativity Research Journal*, *20*(4), 358–364.

Campbell, D. T. (1960). Blind variation and selective retentions in creative thought as in other knowledge processes. *Psychological Review*, *67*(6), 380–400. doi:10.1037/h0040373

Chi, R. P., & Snyder, A. W. (2011). Facilitate Insight by Non-Invasive Brain Stimulation. *PLoS ONE*, *6*(2), e16655. doi:10.1371/journal.pone.0016655

Chronicle, E. P., MacGregor, J. N., & Ormerod, T. C. (2004). What Makes an Insight Problem? The Roles of Heuristics, Goal Conception, and Solution Recoding in Knowledge-Lean Problems. *Journal of Experimental Psychology: Learning, Memory, & Cognition January*, *30*(1), 14–27.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, *2*(10), 406–416.

Davidson, J. E. (2003). Insights about insightful problem solving. In J. E. Davidson & R. J. Sternberg (Eds.), *The Psychology of Problem Solving* (pp. 149–175). Cambridge: Cambridge University Press.

Duncker, K. (1945). *On Problem-Solving* (Vol. 58). Washington: American Psychological Association INC.

Fleck, J. I., & Weisberg, R. W. (2013). Insight versus analysis: Evidence for diverse methods in problem solving. *Journal of Cognitive Psychology*, *25*(4), 436–463. doi:10.1080/20445911.2013.779248

Haider, H., & Rose, M. (2007). How to investigate insight: A proposal. *Methods*, *42*(1), 49–57.

Jones, G. (2003). Testing two cognitive theories of insight. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *29*(5), 1017–1027.

Katona, G. (1940). *Organizing and memorizing: studies in the psychology of learning and teaching*. New York: Columbia University.

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *25*(6), 1534–1555.

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory & Cognition*, *29*(7), 1000–1009.

Knoblich, G., Öllinger, M., & Spivey, M. J. (2005). Tracking the eyes to obtain insight into insight problem solving. In G. D. M. Underwood (Ed.), *Cognitve processes in eye guidance* (pp. 355–375). Oxford, UK: Oxford Univ. Press.

Köhler, W. (1925). *The mentality of apes*. New York: Livewright.

Kounios, J., & Beeman, M. (2014). The Cognitive Neuroscience of Insight. *Annual Review of Psychology*, *65*(1), 71–93.

Lovett, M. C., & Anderson, J. R. (1996). History of success and current context in problem solving: Combined influences on operator selection. *Cognitive Psychology*, *31*(2), 168–217.

Luo, J., & Niki, K. (2003). Function of hippocampus in “insight” of problem solving. *Hippocampus*, *13*(3), 316–323.

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, & Cognition*, *27*(1), 176–201.

Metcalfe, J., & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition*, *15*(3), 238–246.

Ohlsson, S. (1992). Information-processing explanations of insight and related phenomena. In M. Keane & K. Gilhooly (Eds.), *Advances in the psychology of thinking* (pp. 1–44). London: Harvester-Wheatsheaf.

Öllinger, M., Jones, G., Faber, A. H., & Knoblich, G. (2013). Cognitive mechanisms of insight: The role of heuristics and representational change in solving the eight-coin problem. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 931–940.

Öllinger, M., Jones, G., & Knoblich, G. (2008). Investigating the Effect of Mental Set on Insight Problem Solving. *Experimental Psychology*, *55*(4), 269–282.

Öllinger, M., Jones, G., & Knoblich, G. (2013a). Insight and Search in Katona's Five-Square Problem. *Experimental Psychology (formerly Zeitschrift Für Experimentelle Psychologie)*, 1–10. doi:10.1027/1618-3169/a000245

Öllinger, M., Jones, G., & Knoblich, G. (2013b). The dynamics of search, impasse, and representational change provide a coherent explanation of difficulty in the nine-dot problem. *Psychological Research*, 1–10.

Öllinger, M., & Knoblich, G. (2009). Psychological Research on Insight Problem Solving. In H. Atmanspacher & H. Primas (Eds.), *Recasting Reality* (pp. 275–300). Berlin-Heidelberg: Springer.

Ormerod, T. C., MacGregor, J. N., & Chronicle, E. P. (2002). Dynamics and constraints in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *28*(4), 791–799.

Reverberi, C., Toraldo, A., D'Agostini, S., & Skrap, M. (2005). Better without (lateral) frontal cortex? Insight problems solved by frontal patients. *Brain*, *128*, 2882–2890.

Sandkühler, S., & Bhattacharya, J. (2008). Deconstructing Insight: EEG Correlates of Insightful Problem Solving. *Plos One*, *3*(1), 1459–1472.

Seifert, C. M., Meyer, D. E., Davidson, N., Patalano, A. L., & Yaniv, I. (1995). Demystification of cognitive insight: Opportunistic assimilation and the prepared-mind perspective. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 65–124). Cambridge, MA, USA: Mit Press.

Simonton, D. K. (1995). Foresight in insight? A Darwinian answer. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 465–494). Cambridge, MA, USA: Mit Press.

Sternberg, R. J., & Davidson, J. E. (1995). *The nature of insight*. Cambridge, MA: MIT Press.

Wallas, G. (1926). *The art of thought*. New York: Harcourt Brace Jovanovich.

Weisberg, R. W. (1995). Prolegomena to theories of insight in problem solving: A taxonomy of problems. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 157–196). Cambridge, MA: MIT Press.

Wertheimer, M. (1959). *Productive thinking*. New York: Harper.

6 APPENDIX – SOLUTION OF THE CROSS-SHAPE KATONA PROBLEM