Please cite as:

Kefalidou, G. (2017). When immediate interactive Feedback Boosts Optimization Problem Solving: A 'Human-in-the-Loop' approach for solving Capacitated Vehicle Routing Problems. *Computers in Human Behavior*, 73, 110-124, doi: http://dx.doi.org/10.1016/j.chb.2017.03.019

When immediate interactive Feedback Boosts Optimization Problem Solving: A ‘Human-in-the-Loop’ approach for solving Capacitated Vehicle Routing Problems

Genovefa Kefalidou[1]

Lancaster University, United Kingdom

Correspondence concerning this article should be addressed to Genovefa Kefalidou, Human Factors Research Group, Faculty of Engineering, The University of Nottingham, NG7 2RD, Nottingham, UK at all times

Email: Genovefa.Kefalidou@nottingham.ac.uk

[1] Present address: Human Factors Research Group, Faculty of Engineering, The University of Nottingham, NG7 2RD, Nottingham, United Kingdom

When immediate interactive Feedback Boosts Optimization Problem Solving: A ‘Human-in-the-Loop’ approach for solving Capacitated Vehicle Routing Problems

Abstract

In past, feedback in problem solving was found to improve human performance and focused mainly on learning applications. Interactive tools supporting decision-making and general problem-solving processes have long being developed to assist operations but not in optimization problem solving. Optimization problem solving is currently addressed within Operational Research (OR) through computational algorithms that aim to find the best solution in a problem (e.g. routing problem). Limited investigation there is on how computerized interactivity and metacognitive support (e.g. feedback and planning) can support optimization problem solving. This paper reports on human performance on Capacitated Vehicle Routing Problems (CVRPs) using paper-based problems and two different versions of an interactive computerized tool (one version with live explanatory and directive feedback alongside planning (strategy) support; one version without strategy support but with live explanatory feedback). Results suggest that human performance did not change when people were given paper-based post-problem feedback. On the contrary, participants’ performance improved significantly when they used either version of the interactive tool that facilitated both live feedback support. No differences in performance across the two versions were observed. Implications on current theories and design implications for future optimization systems are discussed.

Keywords: human-in-the-loop; interactive route optimization; human performance; concurrent feedback; metacognition

1. Introduction

Enhancing human performance in problem solving through the use of technology-oriented interventions has been the focus of research within the past decades. Programmed instruction aims to provide triggers to the solver or learner to perform and adopt a target behaviour while reinforcing it (Bangert-Drowns, Kuli, Kulik & Morgan, 1991). Despite the breadth and spread of different technology-enhanced innovations to support human problem-solving, evidence for improvements in human performance have been scarce and inconsistent when tested in laboratory settings. For example, it was found that immediate feedback provision could disrupt performance and task execution (see e.g. Munro, Fehling & Towne, 1985; Schmidt, Young, Swinnen & Shapiro, 1989; Schooler & Anderson, 1990) but could also prevent accumulating error-detection skills. Research in learning and problem solving have acknowledged the so-called 'assistance dilemma'.

According to the notion of 'assistance dilemma' there is a debate as to how much assistance and instructions individuals need to have available to complete a problem-solving task (Koedinger & Aleven, 2007) suggesting a need to define a 'golden threshold' for balancing the levels and volume of information and assistance is being provided to users in completing a learning and problem solving task. Research on error-trapping and feedback provision during problem solving, on the other hand, seems to be inconsistent with some researchers claiming that immediate feedback when solving problems is good and can provide good performance (Corbett & Anderson, 2001; Kapa, 2001) and others saying the opposite (Schooler & Anderson, 1990; Cheshire, Ball & Lewis, 2005; Lurie & Swaminathan, 2009).

Timing of feedback has also been found responsible for differences in human performance. For example, interruptive feedback (e.g. concurrent feedback) from an external source (e.g. human tutor or computerized tutor) during solver's or learner's engagement with the task seems to inhibit learning (e.g. Corno & Snow, 1986). Kalhuvy and Anderson (1972) proposed the interference preservation hypothesis according to which concurrent (or immediate feedback) can result to interference caused by error-flagging and the searching for correct response. As such, delayed feedback (or else known as Delay Retention Effect – DRE) is more favourable because errors are expected to be forgotten and not interfere with retention of information for the task completion. A number of studies support delayed feedback (e.g. Kulhavy & Anderson, 1972; Surber & Anderson, 1975) but there are a number of studies that oppose DRE favoring immediate feedback provision (e.g. Corbett & Anderson, 1989, 2001; Dihoff, Brosvic, Epstein, & Cook, 2003). Shute (2008) provides an excellent review on formative feedback issues.

More recent studies in computerized feedback approaches focussed on investigating three major principles of feedback: presence of feedback, timing of feedback and content of feedback. Corbett and Anderson (2001), for example, tested students on computer language programming problems in 4 different tutor feedback conditions: 1) immediate feedback *and immediate error correction (force mechanism to correct the error immediately otherwise students were unable to proceed)*, 2) *immediate error flagging (control mechanism for students for error correction – students do not have to fix the error immediately in order to proceed*), 3) feedback on demand *(control mechanism for students for error correction – students do not receive any feedback unless they request it)* and 4) no computer tutor support (control group – students do not receive any 'symbol-by-symbol' feedback support but they do receive feedback at the end of the programming task on whether the program works correctly). The immediate error correction feedback group (that was exposed to the force mechanism) with greatest tutor control of problem solving yielded the most efficient learning (i.e., the first condition). Corbett and Anderson (2001) posited that immediate corrective feedback provided by computer systems (e.g. intelligent computer tutors) during problem solving resulted in faster time completion. Students in all the computer tutor support conditions performed equally well having more successful learning experience while performance was improved. According to Corbett and Anderson (2001), the time when feedback is given is critical and heavily depends on the nature of the problem-solving tasks.

Kapa (2001) found similar results in a study with 441 teenager pupils when asked to solve mathematical word problems using a computerized learning environment that supported directive feedback during problem solving process and corrective/directive feedback at the end of problem solving process. What she found was that pupils' performance was improved significantly in all conditions where feedback was present during the problem solving process but not when it was present at the end of the problem solving phase. Kapa posited that the reason why directive feedback during the problem solving process was more successful than corrective/directive feedback at the end of the process was because directive feedback allowed pupils to make connections with prior knowledge making the subject matter more memorable and distinctive in their cognitions. On the other hand, pupils that were exposed to corrective/directive feedback at the end of the problem solving process were asked to be more creative (e.g. search for alternative solutions), something that they may not have found necessary (i.e. not motivated enough) as they had already provided a solution to the problem. As a consequence of this 'de-motivation', their performance was decreased compared to the pupils with directive concurrent feedback.

Azevedo and Bernard (1995) conducted an extensive meta-analytic study looking at the effects of computerized feedback on human performance and learning in 22 studies. They found out that the provision of immediate feedback is a critical factor determining computerized instruction and learning and constitutes and instructional advantage. Achievements and outcomes were greater for the experimental groups exposed to feedback compared to control groups that had no feedback provision. Individuals presented with computerized feedback appeared to learn better compared to those that received no computerized feedback. In later research Azevedo and Hadwin (2005) acknowledged the need for adaptive dynamic self-regulatory mechanisms that support scaffolding and planning (e.g. strategic and procedural scaffolds) in learning (Azevedo & Hadwin, 2005). However, more research is necessary to understand the levels and types of metacognitive mechanisms that can improve learning and problem-solving outcomes. For example, which and what procedures and embedded tools within a computerized environment can aid problem solving and learning activities? Also, what metacognitive approaches are and when more beneficial for human performance?

Cheshire, Ball and Lewis (2005) on the other hand, found that when they tested children's performance on analogical reasoning tests across 'feedback', 'no feedback', 'explanation and feedback', 'explanation', 'practice' and 'control' conditions, the best performances were generated when children were provided with 'explanation and feedback' or just 'feedback' while solving the problems (where feedback involved providing children with a responses as to whether their solution was correct or incorrect). This suggests that providing feedback at the end of the problem improved children's' performance facilitating retention while demonstrating a higher feedback value over explanation as feedback on its own (compared to explanation provision on its own) promoted better problem solving performance in children. Despite the differences found in literature in regards to the effects of time on feedback, it seems that presence of feedback (at the end of the process) is beneficial for both quality of human performance and time completion as well supporting a more metacognitive theoretical approach (instead of a mental representation approach) in analogical problem solving that facilitates the generation of strategies and higher-order thinking processes.

Schooler and Anderson (1990) also suggested that delayed feedback -instead of timely feedback - gives an advantage on performance as it triggers less error-prone behaviour by encouraging participants to reflect upon and evaluate their own performance in finding errors; this consequently, had as a result to generate better performance. Similar results were also recently found in dynamic decision-making research (e.g. Lurie & Swaminathan, 2009). Lurie and Swaminathan (2009) found that when real-time feedback was increased on prior decisions, performance declined. Frequent feedback appeared to either distract or even facilitate fixation to recent data impairing the ability to synthesize prior information adequately. It is important to note that these observations are referring to feedback that is not constantly visually provided to the participants throughout the problem-solving and/or learning task.

1.1 The Value of Optimization Problem Solving – The Case of Capacitated Vehicle Routing Problems (CVRPs)

Optimization problem solving constitutes a cognitive process that relates strongly to everyday applied tasks (e.g. finding the best and most efficient route between two locations) and has strong applications in the field of Operational Research (OR) and Transport Logistics (Danzig & Ramser, 1959; Toth & Vigo, 2002). Capacitated Vehicle Routing Problems (CVRPs – see Figure 1 for an example) is an optimization routing problem, where a fleet of trucks is responsible to deliver goods from a depot to a set of customers (under a number of constraints such as weight load allowed per truck and number of routes) are being utilized by Logistics and Supply Chain industries to perform deliveries CVRPs' objective is to minimize the total length of the distance travelled. Use of computational algorithms such as exact algorithms, heuristics and metaheuristics that address large instances (e.g. high number of customers to deliver items to) are currently being employed for solving such problems (Blasum & Hochstaettler, 2000; Fisher, 1994; Lee & Mitchell, 1998; Toth & Vigo, 2001; Burke & Kendall, 2005).

Despite the efficiency computational algorithms (e.g. exact algorithms, heuristics and metaheuristics) that tackle large instances for such problems (Blasum & Hochstaettler, 2000; Fisher, 1994; Lee & Mitchell, 1998; Toth & Vigo, 2001; Burke & Kendall, 2005) appear to exhibit their results are either problem-specific or they are dependent on choosing the most appropriate evaluation functions running the risk of generating sub-optimal or local optimum results.

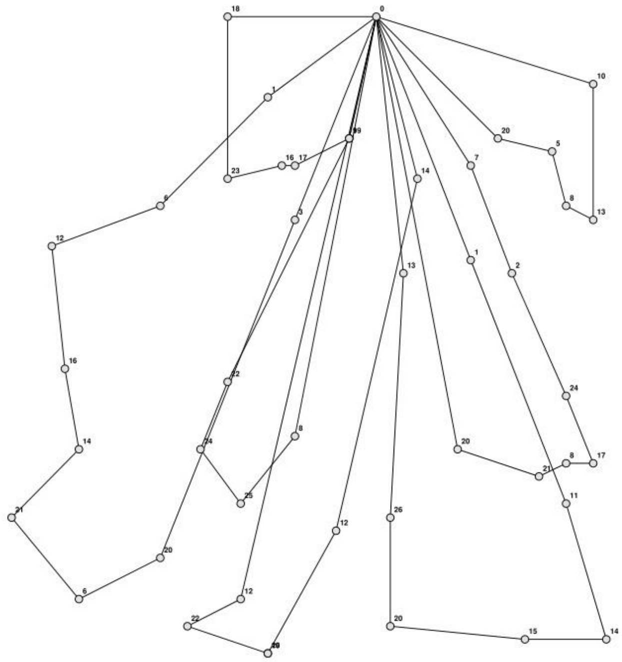

Figure 1. Example of a CVRP with its optimal solution.

Furthermore, there is little evidence that such computational strategies provide optimal solutions in a reasonable time frame as the number of customers increase (Maringer, 2006; Michalewitz & Fogel, 2004; Hu, 1982) making debatable the efficiency they offer (Ioannou, 2008; Crainic, Gendreau & Potvin, 2009). As such, unfortunately, there is not yet a complete algorithm that can provide a solution that is both quick and optimal when the number of customers is large.

Considering the computational weaknesses, previous psychological literature has investigated how humans solve such hard problems (see e.g. MacGregor & Ormerod, 1996 on human performance on Traveling Salesman Problem (TSP)). MacGregor and Ormerod (1996) examined drawn solutions to Traveling Salesman Problem (TSP) instances (a variation of CVRPs with no weight constraint and a single route to draw) and found that, up to 100 nodes, human solutions were comparable with heuristic computer methods acknowledging the potential value that human solutions can have in solving optimization problems. However, such research tested human performance only without speculating (or testing) metacognitive mechanisms that may explain performance levels.

Further attempts in examining how to improve current computational approaches for optimization problem solving focused on developing interactive optimization approaches, the aim of which is to bring the human operator 'in-the-loop'. Examples include solving everyday problems (Boyd & Vandenberghe, 2004) to more specialized ones such as spacecraft design (Fukunaga & Stechert (1997) and routing problems (Waters, 1984; 1987; Sreevalsan et al., 2007; Klau, Lesh, Marks & Mitzenmacher, 2010). What all these interactive optimization attempts offered was either algorithmic real-time reconfiguration, which almost entirely excluded any human involvement in the execution and procedural loop (i.e. human almost out-of-the-loop responsible for refining algorithmic parameters and not constructing routes from scratch) or allowed direct human intervention where human solutions would be 'refined' by computational algorithms. They all reported however, improvements in optimization results as a product of these interactive synergies.

What the above approaches have not explored was the role of feedback provision and its effects in boosting human performance. It seems that there is a great need to account for the use of additional higher order cognitive processes such as meta-cognition (e.g. planning, monitoring and evaluative feedback) to enhance optimization problem solving. For example, error-trapping has not been tested or even employed as a mechanism within computational approaches yet and has not been investigated as a self-regulatory mechanism in human performance in optimization problem solving. The research presented in this paper aims to test the provision of feedback (a metacognitive construct) in CVRPs (an optimization problem) within both paper-based and computerized settings and explore its effect on human performance in solving optimization problems such as CVRPs.

Considering the positive effects that feedback mechanisms appear to exhibit in human performance (e.g. Shute, 2005), the lack of examination of effects of these in route optimization problem solving and the need to re-think the current approaches employed in optimization computational algorithms (i.e. consideration of procedural cognitive processes rather than adoption of metacognitive mechanisms), this paper aims to test the effects that metacognitive approaches such as feedback provision (as a self-regulatory mechanism) on both paper-based and computerized settings have on human performance on optimization problem solving such as CVRPs. As part of the task environment manipulation and considering the current literature on the effects of the timing of feedback provision, the author varied feedback provision timing while considering the presence of additional metacognitive elements such as heuristics planning provision.

Furthermore, while prior research focused on testing 'human-in-the-loop' frameworks in optimisation tasks where computational algorithms were either embedded within the system or operators manipulated solutions, in this paper, the author presents a first step in applying the 'human-in-the-loop' framework in solving CVRPs without the use of any computational algorithm but by providing computational support for certain aspects of the CVRPs (i.e. calculations) and direct route construction by humans. In doing so, the author aims to provide an empirically cognitive-driven approach in direct route planning, providing insights as to how human cognition can enrich this task. The author reports results from a lab-based study that looked upon human performance when interacting with a CVRP computerised tool.

This paper presents the results from two lab-based experiments on solving CVRPs.

In Experiment 1, the investigation on how humans solve CVRPs under a post-problem feedback condition took place. In Experiment 2, an interactive computerized CVRP tool that facilitates a 'Human-in-the-Loop' approach, was adopted to test human performance when solving paper-based and computer-based CVRPs. One version of the tool provided concurrent feedback on violation of problem constraints and the second version of the tool (i.e. full version) provided concurrent feedback on route length (as each route was constructed), feedback on constraints violation and facilitated assistance in route planning through the provision of two types of heuristics (Calculating and Clustering) that prior research (Kefalidou & Ormerod, 2014) found to be predictors of good human performance in CVRPs. Within the context of the presented research, humans played an active role (rather than a supervisory role) in solving CVRPs. Human performance was captured in both experiments by measuring participants' solutions quality (i.e. how far away their solutions were from the best known solutions as published within the current OR literature).

As such, the purpose of this paper is firstly to draw upon principles and applications of feedback as developed within literature in learning and identify how it manifests within a different domain – that of the optimization problem solving literature. Secondly, the present research aims to identify the impact computerized feedback has on optimizing human performance in a route optimization problem following a joint-cognitive approach. It is important to emphasize that this is a first attempt in adopting a joint cognitive system approach in a route optimization problem setting where humans provide solutions without any assistance from a computational algorithm and provides a test-bed for understanding how people optimize (i.e. attempt to find best solutions) through metacognitive feedback and planning provision. While notions of learning (and accumulation of knowledge) is an intrinsic element of human problem solving performance, the aim of this paper is to move a step back and look upon human performance *per se* and not learning or skills acquisition (the author argues that this would be a potential next step to investigate).

1. Experiment 1

In Experiment 1, the aim was to observe any differences in human performance when provided with post-problem feedback. The aim of the error-trapping mechanism was to investigate whether providing post-problem feedback would improve or not human performance. In this condition, if participants produced invalid solutions, they were asked to re-visit the CVRP and find a valid solution before they proceeded further (utilizing a similar control mechanism with prior research (e.g. Corbett & Anderson, 2001 and Chi et al., 1989). However, as the aim of the experiment was to identify effects of post-problem feedback only (and not for example effects of automatic immediate correction) on human performance, the author followed Chi et al's (1989) approach of feedback provision (i.e. if errors are found in solutions after problem completion, then the solutions are returned back to participants for correction) rather than Corbett and Anderson's (2001) approach on providing forced immediate correction during the problem solving process.

Considering prior research on feedback suggesting that delayed feedback can be beneficial for human performance (known as Delay Retention Effect – DRE) as they not interfere with retention of information for the task completion (e.g. Kulhavy & Anderson, 1972; Surber & Anderson, 1975; Cheshire et al., 2005; Schooler & Anderson, 1990), it was hypothesized that post-problem feedback will improve human performance. Furthermore, considering for the posterior timing in the provision of feedback, it was anticipated that this would have an effect on total problem time completion. The fact that provision of feedback was absent in prior research on computerized interactive optimization problem solving (e.g. Waters 1984; 1987; Klau et al., 2010) informed heavily the motivation and rationale for testing effects of feedback provision on optimization human problem solving performance. The two hypotheses are presented below.

H1: Post-problem feedback will improve human performance.

H2: Post-problem feedback will affect time completion.

The control condition of the experiment was to solve the problems without being forced to produce a valid solution.

3.1. Method

3.1.1. Participants

Sixty-nine postgraduate students (35 female and 34 male; mean age = 21.8) from XXXXXXXXXXX University participated (32 in control condition and 37 in the experimental). They were recruited on a credit-based system. Participants were both English-native and foreign speakers. Informed consent was obtained from all participants.

3.1.2. Task and Materials

Participants solved five paper-based stimuli with CVRP nodes drawn on them, taken from prior research e.g. Kefalidou & Ormerod (2014). See for an example in Figure 2 below. Each problem had the depot in a different position (green node on Figure 2 with 0 weight load). Participants drew solutions using different color pencils and a camera recorded participants' hands when drawing. The constraints and guidelines for Experiment 1 were taken from Kefalidou & Ormerod (2014).

Figure 2. CVRP Paper-based Example Stimuli

3.1.3. Design

The experiment had two independent variables. A within-subjects factor which was the CVRP Problem, with four levels (nodes and routes respectively being 33-4, 33-5, 39-6 and 45-7) while the second was a between-subjects factor of Environment with two levels (i.e. error-trapping versus control). Dependent variables were the Percentage Above Optimal (PAO) for each solution attempt, and Solution Time (sec). Optimal solutions were calculated using a Tabu Search algorithm (Mu, Fu, Lysgaard & Eglese, 2009).

3.1.4. Procedure

Each participant was given a paper sheet with the general instructions for the experiment and a printed book containing the task instructions and stimuli. Each problem had its instructions on the same page where the participants had to solve the problem (see Figure 2). Participants were requested to use different colored pencils to design different routes and was explained to them that if they wanted to make further calculations they could use the stimuli paper where the problems were displayed on (see Figure 2). The first problem that they were given to solve was a practice problem and as such it was not included in the analysis phase.

The CVRPs were given to each participant in a random order to minimize practice effects. Participants could ask any questions –if they needed- throughout the duration of the experiment. There was no time restriction, however, participants were made aware that the experiment takes approximately 1 hour.

In the control group, participants were not exposed to 'error-trapping' and post-problem feedback mechanisms and as such once they finished solving each problem, they immediately moved on to solve the next problem in line. On the contrary, participants in the experimental (i.e. 'error-trapping' and post-problem feedback) group, were exposed to post-problem feedback at the end of the completion of each problem. If their solutions were found with errors, then they were requested to revisit and correct their solutions prior to proceeding to the next problem. They were not allowed to proceed to the next problem unless they had corrected the current erroneous solution.

1.1. Results

Each solution was checked for calculation errors and/or violations of problem constraints. Solutions containing any errors (from the control group) were disregarded from the quantitative analysis while erroneous solutions were tracked as part of the error-trapping mechanism and the post-problem feedback in the experimental condition (see table X below for more details). Similarly to XXXXXXXXXXXXXXXXXXXXXXXXX and XXXXXXXXXXXXXXXXXXXXXXX, each participant's solution routes were measured with the use of software to calculate the individual route length and total solution length.

Table 1.

Errors occurred in Experiment 1 across conditions

	Error	No of Participants
Control	Weight constraint	3
	Node exclusion	1
	Route violation	1
Experimental (Error-trapping)	Weight constraint	10
	Node exclusion	1
	Both	1

Note: the Error-trapping group errors were not excluded from the analyses; instead participants were required to revise their solution and generate a valid solution. Their solutions were included in the analyses.

As Table 1 suggests, participants in the Experimental condition (error-trapping and post-problem feedback) performed approximately double number of errors compared to the control group. It is not clear the reason why this might have happened and further research is needed to ascertain this.

1.1.a. Percentage above optimal (PAO) data

Mean PAOs are shown in Table 2. The error-trap group performed better than the control group for the 33-4 problem (mean PAO error-trap = 6.56 vs. mean PAO control = 7.31) and the 33-5 problem (mean PAO error-trap = 12.18, control = 14.20). The pattern was reversed for the 39-6 and 45-7 problems (39-6: mean PAO control = 16.43, error-trap = 20.52; 45-7: control = 16.0, error-trap = 17.51).

Table 2:

Mean % above optimal (PAO; standard deviations in brackets) for each problem in error-trap and control groups.

	P2 (33-4)	P3 (33-5)	P4 (39-6)	P5 (45-7)
Error trap	6.56*	12.18	20.52	17.51
	(5.57)	(10.77)	(15.40)	(7.35)
Control	7.31*	14.20	16.43	16.06
	(8.71)	(13.29)	(14.30)	(7.17)
Overall Mean	6.88*	13.06	18.73	16.88
	(7.05)	(11.87)	(14.95)	(7.25)

Note: * indicates closer to optimal results (where optimal is 0)

A 2 (Group) x 4 (Problem Complexity) ANOVA on the Percentage Above Optimal (PAO) scores revealed a significant effect of problem $F(3,180) = 26.46$, $\eta_p^2 = .306$, $p < .001$. There was no significant interaction between Problem and Environment, $F(3,180) = 1.82$, $\eta_p^2 = .029$, $p = .145$. No significant effect of Environment was found, $F(1,60) = .101$, $\eta_p^2 = .002$, $p = .752$. Pair-wise comparisons were conducted and revealed that PAO scores on 33-4 and 33-5 problems were significantly lower (i.e. more optimal) than the PAO scores on 39-6 and 45-7 problems, $p < 0.05$, and there was no difference on PAO score between 39- 6 and 45-7 problems, $p = 1.00$. This suggests that participants performed better in problems with lower number of nodes and routes.

As such, results suggest that H1 (Post-problem feedback will improve human performance) is not supported.

1.1.b. Solution times

A two-way ANOVA on solution times revealed a significant effect of Environment, $F(1,62) = 5.64$, $\eta_p^2 = .061$, $p < .05$ but a non-significant effect of time completion of problems, $F(3, 186) = 2.07$, $\eta_p^2 = .032$, $p = .105$. Most importantly, there was no significant interaction between time completion of problems and Environment, $F(3, 186) = .990$, $\eta_p^2 = .016$, $p = .398$ suggesting that error-trapping mechanism had no effect on time completion. Table 3 shows the means and standard deviations of the time completion data. As can be seen, the control Group performed faster than the error trap group in all conditions.

Table 3.:

Mean Solution Times (sec), (Stds in brackets) for each problem in 'error trap' and control' groups.

	P2 (33-4)	P3 (33-5)	P4 (39-6)	P5 (45-7)
Error trap group	573* (259.53)	638 (1047.52)	597 (224.01)	859 (418.28)
Control group	470* (189.55)	515 (260.72)	477 (171.36)	529 (277.68)
Overall Mean Time Completion	528* (235.57)	584 (801.60)	544 (209.86)	715 (396.97)

Note: * indicates faster results

Considering the Time Completion results, it appears that H2 (Post-problem feedback will affect time completion) was not supported.

1.2. Discussion

Overall, results from Experiment 1 confirm prior results in this field (e.g. Kefalidou & Ormerod (2014)), in that the participants performed worse on the same problems in both experiments (P4 - 39-6) even though prior research focussed on different experimental factors (e.g. verbalization and cognitive load) rather than error-trapping and post-problem feedback, which was the experimental factor of this lab-based study. A significant effect of problem complexity was found, indicating that the problems with higher numbers of nodes and routes to draw (e.g. 39-6 and 45-7) were seemingly considered more difficult to solve as people generated less optimal compared to less complex problems (e.g. 33-4 and 33-5). This finding was expected as the complexity of a problem suggests its difficulty and necessitates more effort in order to be successfully solved.

Although effects of error trapping appear to vary by problem, none of the effects on time completion were significant. Furthermore, there was no interaction between problem complexity and environment (i.e. error-trapping and post-problem feedback). This implies that providing post-problem feedback to the solvers does not have any impact on their performance contrary to H1 (i.e. Post-problem feedback will improve human performance). This currently appears not to support prior research on positive effects of post-problem (or delayed) feedback on human performance (see e.g. Kulhavy & Anderson, 1972; Surber & Anderson, 1975; Cheshire et al., 2005).

. A suggestion to explain the occurrence of no effects, might be that the timing of feedback is a factor that may influence human performance on CVRPs. For example, if feedback was given *during* the problem-solving process (instead of being given at the *end* of the problem solution), human performance might have been improved. As discussed earlier in the literature, there is a body of research that supports the notion that concurrent immediate feedback provision can be beneficial to human performance in learning and problem solving (see e.g. Corbett & Anderson, 2001; Kapa, 2001; Azevedo & Bernard, 1995). This is further investigated in Experiment 2. Such an improvement might occur because immediate and concurrent feedback allows humans to re-plan and re-evaluate their solving behaviour (as their problem space is dynamically changed), allowing them to minimize the chance of getting trapped in the same errors. Furthermore, concurrent live feedback promotes self-awareness and conscious processing in problem-solving. As it has already been discussed earlier, there is research that accounts for such a possibility (Corbett & Anderson, 2001; Cheshire, Ball & Lewis, 2005).

Concerning time completion, there was a non-significant effect of time completion and no interaction between error-trapping and time completion suggesting that the effect on time is the same regardless of what condition (error-trapping or no error-trapping) someone is in. Also, completion time and PAO were not positively correlated, $r = -.013$, $n = 70$, $p = .285$. However, as the number of nodes and routes to design increased (problems P4 (39-6) and P5 (45-7)), the control group performed better than the error trap group.

Although this was not an a priori prediction, one interpretation is that for the more complex problems (i.e. P4 (39-6) and P5 (45-7)) , the error trap group needed more time to solve them as the searching process for a valid solution was increasing, something that was reflected in their decreasing performance. Their decreasing quality of performance may be taken as an indicator of a stronger 'satisficing' attitude as their aim was to find an acceptable solution. This comes in alignment with Simon's position (Newell & Simon, 1956) where solvers perform a serial search of item-by-item and they finish it as soon as they find a 'satisficing' utility. It seems plausible that when the solvers from the 'error-trapping' group were facing more complex problems, their performance was getting worse, yet valid and 'satisficing' because they did not put further effort to produce a better (or even optimal) solution. However, in problems with lower complexity (e.g. 33-4 and 33-5), the same participants performed better than the participants in the control group. A possible explanation for this observation is that the error-trap participants perceived and tackled those problems with a more 'optimizing' attitude due to lower problem complexity. This is an avenue for further future research. On a similar note, it would be expected that post-problem feedback and error trapping group should perform better in all problems. However, looking more carefully the problem-by-problem performance, there seems to be no linearity. This may be taken to suggest that learning does not seem to occur within the timeframe of this study as performance is not steadily increasing. Similarly, demotivation (as an explanation for decreased performance) may not be exhibited for the same exact reasons. However, more research is needed to ascertain this speculation.

2. Experiment 2

Considering the results from Experiment 1 in regards to post-problem feedback and error-trapping, this second experiment investigated how humans solve CVRPs when the task environment changes in regards to visual and interactive aspects. Particular interest was whether human performance would be improved via the provision of concurrent and immediate feedback in this instance. The utilization of the computer-based tool to solve CVRPs was designed to provide concurrent live feedback to solvers in regards to the quality of their solutions driven by previous research literature on computerized feedback (e.g. Corbett & Anderson, 2001; Kapa, 2001; Azevedo & Bernard, 1995) that suggest that concurrent immediate feedback improves human performance in problem solving and learning. Furthermore, time completion results in Experiment 1 suggested that post-problem feedback did not influence human performance in terms of quality of solutions and as such it is further explored within this Experiment 2 to investigate whether time completion results will be confirmed in the case of concurrent metacognitive support provision.

While in previous research in this field (e.g. see Kefalidou & Ormerod, 2014), task environment was manipulated in terms of calculation demands, in the current experiment, the manipulation was three-fold. Firstly, the stimuli upon which CVRPs were presented was computer-based rather than paper-based, aiming to test human performance when interacting with a different task environment. Secondly, certain aspects of the problem task such as calculations were supported directly by the computer-based environment (through the provision of directive and explanatory concurrent feedback support), while planning support was offered as well (e.g. heuristics provision) based on prior research suggesting that two strategies (i.e. Calculating and Clustering) in CVRPs problem solving appear to predict good human performance – see Kefalidou & Ormerod (2014)) aiming to test whether concurrent (as opposed to Experiment 1 post-problem feedback) would have an effect on human performance. Thirdly, the interactive computerized support tool (from now on it will be referred to as the CVRP Problem Solver – CVRPS) was implemented with two modes: one constituted the full mode of the computer system (incorporating live user feedback on length of routes and weight values (i.e. concurrent directive feedback) and constraint violations (i.e. concurrent explanatory feedback) as well as planning support (through heuristics provision) and the second was a stripped version of CVRPS (incorporating no concurrent live feedback for routes length and weight values (but for constraints violations i.e. concurrent explanatory feedback) and no planning heuristic support. Heuristic and planning support was formulated based on prior research where human heuristics predicting optimal performance were found (e.g. calculating and clustering - see e.g. Kefalidou & Ormerod, 2014) and aimed to provide an additional metacognitive mechanism to solvers (alongside with feedback provision) that assists them in planning their route solutions.

More specifically, the second mode of CVRPS integrated some features only of the full CVRPS providing a computerized environment of the paper-based version. The only difference it had from the paper-based version was that it was able to generate discrete messages to the solvers (i.e. concurrent live feedback) in case they were *violating a constraint of the CVRP* (e.g. missing out a node from the solution or violating the weight constraint per route) only and did not provide feedback in regards to the quality of their solutions. The reason facilitating this kind of feedback within this version of CVRPS was to ensure that participants did not violate the basic constraints of CVRPs and for testing whether concurrent explanatory live feedback provision had any effect on their performance (as compared to Experiment 1).

Prior work on interactive tool-based optimization (e.g. Water's (1984) Klau et al. (2010) work) did not include investigations on whether systems that provide explanatory and discreet concurrent feedback could potentially evoke a better human performance on solving optimization problems. No investigations were either performed to test how human optimization problem-solving develops when interacting with a computer-based tool – instead, prior research focussed mainly on tracking Human Machine Interactions (HMIs) when computational algorithms had already initiated the solutions (e.g. Waters (1984) or on utilizing computational algorithms and human operator interventions as 'refinement' mechanisms of generated solutions (e.g. Klau et al., 2010). . People approach problems differently when they have to start solving them from scratch rather than completing a solution. Starting a problem from scratch gives to the solver the opportunity to develop their own understanding of the problem and model mentally their own pool of potential solutions. On the contrary, if a human solver approaches a problem without them initiating the solution, then they have an additional cognitive load to 1) understand a problem, 2) pick-up from where any machine solutions left and 3) decide upon what remaining course of action to follow, which arguable increases the task complexity. In this experiment, the quality of humans' performance when using interactive computational tools was compared to human performance in paper-based settings.

According to prior research (e.g. Corbett & Anderson, 2001; Kapa, 2001), immediate corrective feedback provided by computer systems (e.g. intelligent computer tutors) during problem solving learning tasks (e.g. learning a programming language and mathematical word problems) and problem solving tasks (e.g. solving analogy problems) resulted in faster time completion and better performance. More specifically, Corbett and Anderson (2001) found before that immediate corrective feedback improves the speed of task completion while computerized feedback improved human performance in learning. Furthermore, research from Kapa (2001) on mathematical word problem solving suggested that metacognitive support such as feedback provision during problem solving improved human performance compared to post-problem support conditions.

As such, it was expected that the use of CVRPS will improve human performance (in terms of both quality and time completion – see Corbett & Anderson (2001) and Kapa (2001) prior research) as it provides metacognitive support such as explanatory and directive concurrent feedback as well as planning support (e.g. Calculating and Clustering heuristics). Participants that are solving the problems using both versions of the software are expected to perform better than the participants that use the paper-based version due to the metacognitive support CVRPS provided to solvers – this included the following support: calculation tasks automatically including the calculation of 1) current total weight and 2) distance covered/route length (via concurrent directive live feedback in full software version) and 3) warning messages provision (via concurrent explanatory live feedback) if a violation occurs (both versions of the software).

It was expected that participants would perform better using the full version of the software not only because of the planning/heuristics support provision but also because it provided two types of feedback provision (i.e. of the automatic calculations performed by the system (as part of the directive concurrent feedback provision) and the warning messages (as part of the explanatory concurrent feedback provision). In the full CVRPS version, participants were provided with the currently total distance covered per route, which was continuously updated every time they made a new action (live directive feedback). This feature was hypothesized that would encourage and help the participants construct short routes and enable them to compare easier alternative routes as per prior research (e.g. Corbett & Anderson, 2001; Kapa, 2001; Azevedo & Bernard, 1995) as it would allow 'on-the-fly' restructuring of their solutions facilitating early error correction and full awareness of their cognitive processes. Finally, the use of CVRPS tool (that facilitates concurrent feedback provision) will have an impact on time completion. Experiment 1 results on time completion suggested that there was no significant difference despite the presence of post-problem feedback support. However, prior research on computerized feedback provision (e.g. Corbett & Anderson, 2001) suggested that immediate corrective feedback resulted to faster time completion rates on the problem solving learning task. Consequently, this begged the question whether the provision of concurrent computerized feedback would have an effect on time completion.

As such four hypotheses were formed to test human performance using CVRPS under these conditions. These are presented below:

H3: Participants' performance will improve when using CVRPS tool

H4: Concurrent live feedback will improve human performance when using CVRPS

H5: Human performance is improved when participants receive both directive and explanatory feedback alongside with planning support (i.e. heuristics support) during the problem solving process using CVRPS tool

H6: The use of CVRPS tool will have an effect on time completion

Therefore, in the current experimental setting, it would be expected that best performances would be observed when using the full version of CVRPS (e.g. as a means for metacognitive support provision through concurrent explanatory and directive feedback and through planning/heuristics support). Furthermore, the worst performances would be expected when participants use the paper-based version of the experiment (facilitating no metacognitive support such as feedback and/or planning support provision).

4.1. Method

4.1.1. Participants

Seventy-seven postgraduate students from XXXXXXXXXXX University participated, 26 in the paper-based condition (13 Female; 13 Male), 26 in the stripped CVRPS version condition (13 Female; 13 Male) and 24 in the full CVRPS version condition. All participants were paid £6 to participate. In the full software version condition there were 11 female and 13 male participants. In total, the number of the female participants was 37 and the number of male participants was 39 (mean age = 21.9). The participants that were recruited were both English-native and foreign speakers. Similarly to Experiment 1, informed consent was obtained from all participants.

4.1.2. Task and Materials

Participants in the two computer-based conditions solved five computer-based stimuli with CVRP nodes drawn on them. The CVRP problems used in this experiment were the same with the ones used in Experiment 1. Problems were displayed using the computer display as shown in Figure 3. Participants in the Paper condition solved the same five CVRP presented in paper-based form (see Figure 2 in Experiment 1). In the paper-based version of the experiment, participants drew solutions using different color pencils to draw the routes and a camera recorded participants' hands when drawing. In the computer-based versions, participants used the mouse to drag-and-drop the pointer to construct each route. The colouring of each route was programmed to be implemented automatically. Participants using the full version of the software were given the opportunity to employ (if they wished to) a ‘Clustering’ strategy through the use of the ‘Group’ option tab of the software and a ‘Calculating’ strategy through the use of the ‘Routes’ option. This formed the metacognitive ‘planning support’ for solving CVRPs. Camtasia, a desktop real-time recording software, was used to log in video files all participants’ activities in the computer-based conditions while they were solving CVRPs. The software also recorded any informal verbalizations made by participants.

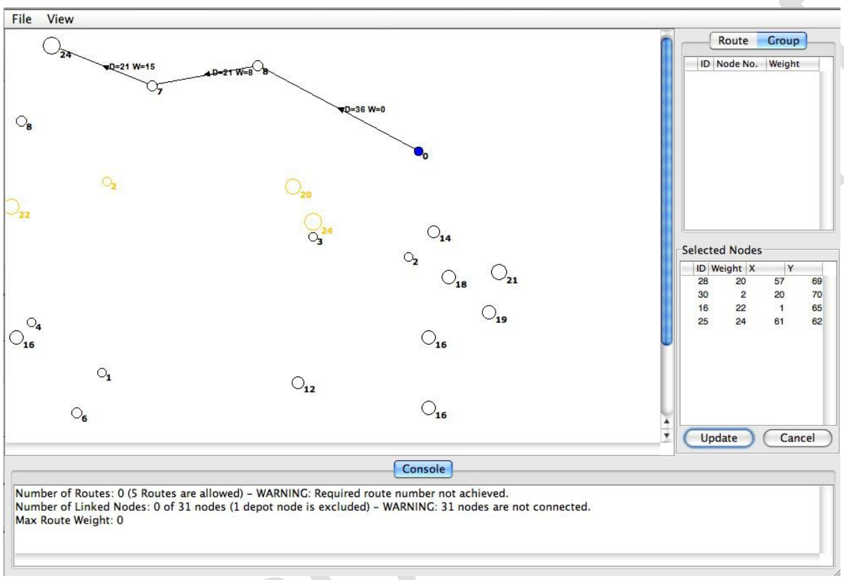

Figure 3. Using the ‘Group’ (Clustering) tab option in the full CVRPS version

The constraints and guidelines for Experiment 2 were the same as in Experiment 1. Although in the paper-based version all the constraints were explained orally to each participant before and during the experiment, in both versions of CVRPS, the participants were provided with two types of metacognitive feedback support: 1) concurrent directive feedback (i.e. ‘live’ information on the route length and node weights as the route was being constructed – see main panel in Figures 3, 4 and 5) and 2) concurrent explanatory feedback (i.e. ‘live’ information on constraint violations – see lower panel in

Figures 3, 4 and 5). Figure 4 shows the messages (concurrent live explanatory feedback and concurrent 'live' directive feedback) and planning (heuristics) provision within the full version of CVRPS. The stripped version of CVRPS provided *only one* type of metacognitive support and that was the concurrent 'live' explanatory feedback on constraints violations (lower panel in Figure 6).

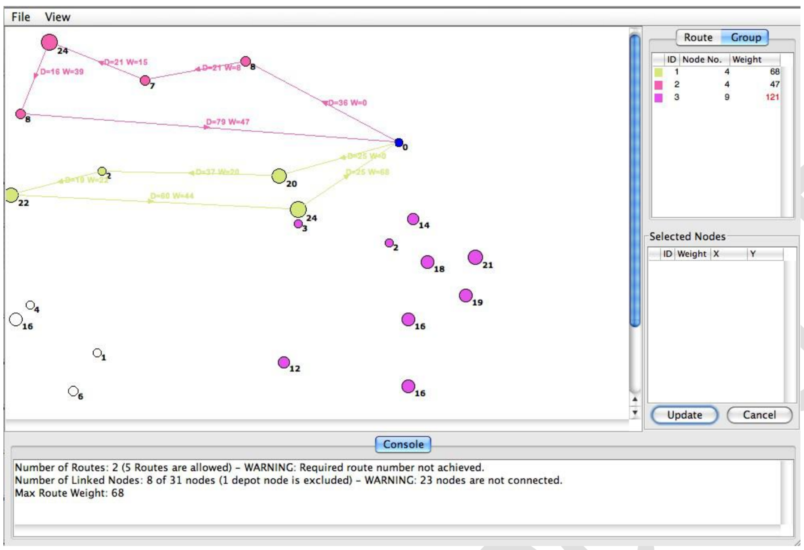

Figure 4. Violating the weight constraint in the full CVRPS version

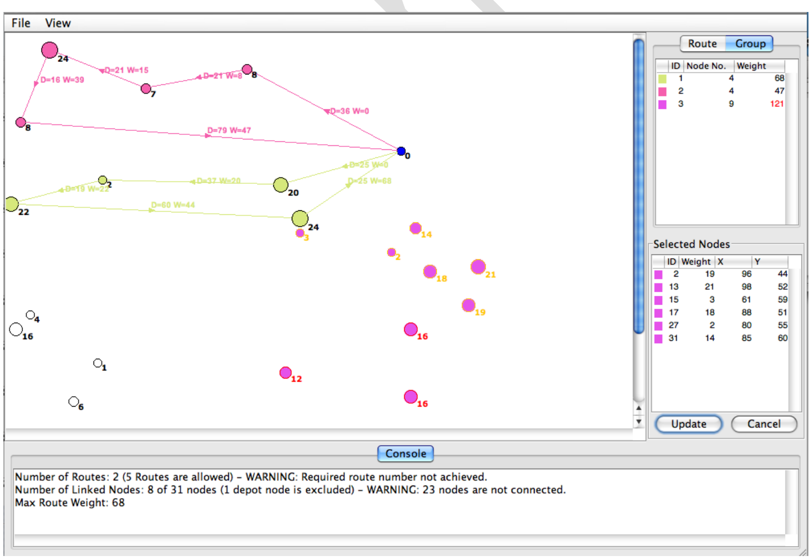

Figure 5. Rearranging cluster of nodes (participants using the planning (heuristics) support in the full CVRPS version)

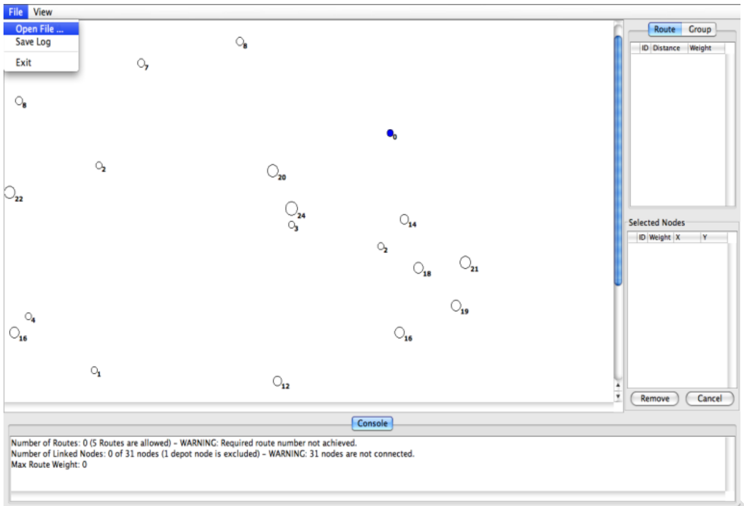

Figure 6. Concurrent 'live' explanatory feedback provision in the stripped version of CVRPS (metacognitive planning (heuristics) support and concurrent 'live' directive feedback was not activated in this version of CVRPS).

4.1.3. Design

The experiment had two independent variables. One was within-subjects: Problem, with four levels (33-4, 33-5, 39-6 and 47-5). The other was a between-subjects factor with three levels (software with strategy support and concurrent live feedback, software without strategy support and with concurrent live feedback, and paper-based with no support). In CVRPS with strategy support and concurrent live feedback, participants were able to access the right-hand panel of the CVRPS where information about each route was displayed (i.e. the total weight and the total distance per route). Also, information on each constructing route was available to them throughout the completion of all CVRPs. In the CVRPS without strategy support (but with concurrent live feedback) condition, participants could neither view the right-hand panel with the route information nor information on constructing each route. What they could view is only a warning message at the bottom of the window in case they violated a problem constraint. In the paper-based version (where no support was provided), the information about CVRPs was presented in an identical manner as in the Experiment 1. Similarly to Experiment 1, dependent variables were the Percentage Above Optimal (PAO) for each solution attempt, and the solution time (sec).

4.1.4. Procedure

Each participant was given a paper sheet with the general instructions for the experiment (see Experiment 1). For the paper-based condition, a printed book containing the task instructions and stimuli was given (see Figure 2 in Experiment 1). Each problem had its instructions on the same page where the participants had to solve the problem. Participants used different colored pencils to design different routes and it was explained to them that if they wanted to make further calculations they could use the paper where the problems were displayed. The first problem was a practice problem and was not included in the analysis.

For the computer-based versions (strategy/concurrent live feedback (directive and explanatory)-support and no strategy-support (with concurrent live explanatory feedback), the researcher initiated CVRPS and loaded each problem to be solved from a text file. An introduction to the use of CVRPS was given to participants before they started the experiment that lasted approximately 10 minutes. Then the researcher initiated the recording software (i.e. Camtasia). Participants then loaded the practice problem. Every time participants were finishing a CVRP problem, they informed the researcher who saved a log text file of the solution of the problem and took a screenshot of the solution. What followed next was the initialization of the next problem.

Similarly to Experiment 1, the control group was not exposed to 'error-trapping' and post-problem feedback mechanisms and as such once they finished solving each problem, they immediately moved on to solve the next problem in line. On the contrary, participants in the experimental (i.e. CVRPS versions) groups, were exposed to concurrent feedback (explanatory and directive) and planning (strategy) support during the problem solving process but rather the feedback was provided as a warning and self-awareness mechanism. As such, they were not 'forced' by the system to provide correct solutions prior to proceeding to the next problem to solve. Participants were able to proceed to the next problem as soon as they provided a solution (valid or invalid) to the current problem.

The problems were presented in a random order in all conditions. Also, it was made clear to all participants that they could ask any questions whenever they felt necessary. There was no time restriction, but participants were told that the experiment was expected to take approximately 1 hour.

4.2. Results

Similar to the previous experiment, each solution was checked for calculation errors or violations of problem constraints. Solutions containing any errors were disregarded from the analysis (these were relevant for the control group only and are the same as the ones presented in Experiment 1 – Table 1). It is important to note here that there was no force mechanism (as prior research employed e.g. Corbett & Anderson, 2001 or similar to Experiment 1 paper-based error-trapping mechanism) to demand or prevent participants from proceeding to the next problem if the solution they provided was invalid. The reason to do so, was that the purpose of this experiment was not to simulate the post-problem feedback condition tested before in Experiment 1 but rather was to investigate potential effects of concurrent feedback provision and planning (heuristics) provision on human performance. Participant solutions in both experimental groups did not have any errors. As in Experiment 1, each participant's solution routes from the paper-based condition were measured with the use of software to calculate the individual route length and total solution length. Participants' solution routes from both the computer-based conditions were provided by CVRPS.

4.2.1. Problem Complexity and Group Effects on PAO

Mean PAOs for all the conditions are shown in Table 4. As it can be seen, the strategy/concurrent live feedback-support group performed better than the other groups for all the problems.

Table 4.:

Mean % above optimal (PAO: Stds in brackets) for each problem in each condition of Experiment 2.

	P2 (33-4)	P3 (33-5)	P4 (39-6)	P5 (45-7)
Planning (Strategy)/concurrent live explanatory and directive feedback support	4.64* (4.57)	8.85 (5.71)	11.34 (7.08)	12.46 (4.37)
No planning/strategy support (with concurrent live explanatory feedback)	5.55* (5.64)	8.92 (7.27)	14.61 (9.35)	13.14 (5.06)
Paper-based (no planning and feedback support – control group)	9.81* (7.79)	13.95 (9.27)	20.51 (15.33)	17.45 (7.80)
Overall Mean PAO	6.58* (6.42)	10.49 (7.78)	15.35 (11.50)	14.26 (6.19)

Note: * indicates closer to optimal results (where 0 is the optimal)

A 2 (Group) x 4 (Problem Complexity) ANOVA on the PAO scores revealed a significant effect of Problem $F(3, 219) = 35.00$, $\eta_p^2 = .324$, $p < .001$. Follow-up pairwise comparisons revealed that PAO scores on 39-6 and 45-7 problems were significantly higher (thus less optimal) than the PAO scores on

33-4 and 33-5 problems – as observed in Experiment 1 as well- , $p < 0.001$ in all comparisons, and there was no difference on PAO scores between 39-6 and 45-7 problems, $p = 1.00$. This suggests that participants' performed better in problems with lower number of nodes and routes.

There was a significant effect of Group, $F(2,73) = 7.09$, $\eta_p^2 = .163$, $p = .002$, planning (strategy)/concurrent live explanatory and directive feedback support : M = 9.32, no planning (strategy)/with concurrent live explanatory feedback support: M = 10.56, paper-based/no support: M = 15.43). Follow-up pair-wise comparisons indicated that participants in the paper-based group have lower PAO scores than participants in the other two groups, planning (strategy)/concurrent live explanatory and directive feedback support vs. paper-based/ no support: $p = .002$; no planning (strategy)/with concurrent live explanatory feedback support vs. paper-based: $p = .012$. There was no significant interaction found between Problem and Group, $F(6, 219) = .819$, $\eta_p^2 = .022$, $p = .556$.

As such, results appear to provide support for H3 (Participants' performance will improve when using CVRPS tool) and H4 (Concurrent live feedback will improve human performance when using CVRPS). However, results suggest that H5 (Human performance is improved when participants receive both directive and explanatory feedback alongside with planning support (i.e. heuristics support) during the problem solving process using CVRPS tool) is not supported as no difference on human performance was observed within the two CVRPS groups.

4.2.2. Solution times

Table 5 shows the means and standard deviations of the time completion (TC) data. A two-way ANOVA revealed a significant effect of problem TC, $F(3, 219) = 9.37$, $\eta_p^2 = .114$, $p < .001$. There was a significant interaction between problem TC and Group, $F(6,219) = 4.34$, $\eta_p^2 = .106$, $p < .001$. There was a non-significant main effect of Group, $F(2,73) = 2.49$, $\eta_p^2 = .064$, $p = .090$.

Table 5:

Mean solution times (standard deviations in brackets) for each problem.

	P2 (33-4)	P3 (33-5)	P4 (39-6)	P5 (45-7)
Planning Strategy /concurrent live explanatory and directive feedback support	508 (195)	420* (292)	582 (307)	651 (298)
No planning strategy support (with concurrent live explanatory feedback support)	598 (230)	482* (246)	582 (246)	835 (409)
Paper-based (no support)	574 (243)	521 (240)	451* (163)	497 (244)
Overall Mean	559 (223)	473* (261)	540 (252)	665 (350)

Note: * indicates faster results

As shown in Table 4, both computer-based groups performed faster with 33-4 and 33-5 problems than the paper-based group, but the reverse effect was found with 39-6 and 45-7. Concerning the speed of the participants when using the software versions, participants in the 'planning strategy/concurrent live explanatory and directive feedback support' Group completed the problems faster than the participants in the 'no planning strategy/with concurrent live explanatory feedback support' Group. This pattern was observed in all problems (mean TC ‘planning strategy/concurrent live explanatory and directive feedback support’ P2 = 508, P3 = 420, P4 = P4 = 582, P5 = 651 vs. mean TC ‘no planning strategy support/with concurrent live explanatory feedback’ P2 = 598, P3 = 492, version’ P4 = 582, P5 = 835). Overall, the fastest performance across all conditions was in P2 problem (33-4) (mean TC = 473) and slowest in the P5 problem (45-7) (mean TC = 665).

As such, H6 (The use of CVRPS tool will have an effect on time completion) was supported.

3. Discussion

The results from this experiment indicate that use of CVRPS to solve CVRP problems improved participants’ performance as they produced even closer to optimal solutions than their fellows in the paper-based version of the experiment. This occurs independently of whether the participants used the full version of the software or just the stripped version. This provides support to the notion previously discussed in literature that concurrent feedback provision improves human performance in terms of response quality and time completion. For example, similarly to Corbett & Anderson (2001) computerized metacognitive support such as concurrent feedback provision improved and sped up human performance suggesting that such metacognitive support effects appear to be transient to other domains of problem solving (i.e. optimization problem solving such as CVRPs). However, as noted earlier in the literature, Corbett & Anderson (2001) found that immediate corrective feedback sped up the time completion process, however, in this research it was found that immediate (i.e. concurrent) directive and explanatory feedback sped up the time completion process. However, it was also found that there are differences in the way that human performance is improved across the versions of CVRPS. For example, results from Experiment 2 suggest that while participants performed best when using CVRPS (which supported metacognitive mechanisms such as feedback provision and planning compared to paper-based version that provided no support) and were faster, there was no significant difference in performance between both CVRPS experimental groups (i.e. concurrent live explanatory feedback support vs. concurrent live directive and explanatory feedback support with planning (strategy) support). This suggests that the provision of additional metacognitive support constructs (i.e. concurrent live directive feedback and planning (strategy) support provision did not improve human performance further. Instead, results suggest that these additional metacognitive supports may have slowed down performance.

The fact that the use of CVRPS (both versions) improved performance compared to the paper-based group (control group with no metacognitive support) but the optimal performances across the two computerized CVRPS versions were not different can be taken to suggest that human performance improvement may have occurred due to the presence of the concurrent live explanatory feedback support provision rather than the presence of planning (strategy) support provision and the presence of concurrent live directive feedback support provision. Instead, it seems as if the presence of planning (strategy) support and the concurrent live directive feedback support only improved time completion and not the quality of generated solutions. However, it is important to note that this has not been observed universally across all problems. This may be due to differentiations on problem complexity (or even cognitive load levels), which is something that future research should explore further.

By providing metacognitive support in CVRPS (e.g. through the provision of two different feedback types and the provision of planning (strategy) support) , participants may have had the opportunity to focus on more 'optimizing' tasks (i.e. focus on quality of solutions rather than focussing on task environment characteristics and procedural tasks such as constraint violations). This is something that participants in the 'paper-based' control Group were not able to do as metacognitive support was offered to them, forcing them to attend to both procedural and 'optimizing' (quality) tasks of the problems, dividing their attention and increasing their task demands. On the contrary, both versions of

CVRPS appear to have supported participants in 'Optimizing' tasks (e.g. more confident readjustment of routes to aim for the shortest one), while off-loading from the task of performing themselves e.g. arithmetic monitoring to ensure that no arithmetic constraints were violated.

Furthermore, the warnings (i.e. concurrent live explanatory feedback) that were provided in both versions of the software (full version and stripped version) improved participants' performance as it was observed within the results in the stripped software version as well. The CVRPS stripped version did not provide participants with any information on routes' length or weight information to assist and support them (no concurrent live directive feedback and planning (strategy) support). Even though this information was missing from the stripped version, participants performed better than in the paper-based version. Looking at the differences in participants' time completion who used both versions of CVRPS (full vs. stripped) and the paper-based version, it was observed that participants who used the full CVRPS version performed faster than the paper-based suggesting that a portfolio of metacognitive support (i.e. two different types of concurrent live feedback and planning support) provides the potential to speed up (but not further optimize) performance. This was expected as participants, in this case, did not need to spend any time in performing calculation tasks or evaluating the length of each route they constructed. This finding is aligned with the time completion results from the stripped version of CVRPS where participants in this instance performed slower than those ones using the full CVRPS version.

Secondly, the explicit planning (strategy) support provision in CVRPS (i.e. Calculating and Clustering) seems to have facilitated faster construction of routes that enabled them to exhibit closer to optimal solutions compared to the paper-based control group. In the stripped version, participants were neither updated for the current total weight or the length of the routes they drew nor was there a suggestion to use the planning support features of the 'full CVRPS version'. Therefore, they needed to spend more time (compared to the full CVRPS version) to perform calculations and evaluate the length of each route by eye. Despite them having to perform these tasks by themselves, the quality of their performance was retained (and in some problem-by-problem cases was better than those in the full version of CVRPS). An explanation for that may be that the warning messages and the concurrent live feedback provided to participants to prevent them from making errors acted as a reinforcement to focus better their attention on the problem tasks reducing the cognitive load.

Regarding time completion performance on the paper-based version, it was observed that participants in this group were faster, yet the worst performers, especially in solving more complex (and larger) problems (i.e. P4 – 39-6 and P5 – 45-7). Although they did not use CVRPS to solve the problems, -thus they needed to perform all the calculation tasks by themselves with no feedback and planning support- they completed the problems faster than participants who used both versions of CVRPS. An explanation to this phenomenon is that for the problems with more demands, features and task requirements, paper-based problems seem to facilitate the problem complexity levels responsible for impairing human performance – these include both calculation and visuo-spatial demands. Instead, the short completion time shown on P4 and P5 problems of the ‘paper-based version’ control Group indicated that participants in this group exhibited a rather ‘satisficing’ approach as they provided not optimal, yet valid solutions faster than their fellows in the other conditions. Another perhaps explanation is that the use of CVRPS involved different ‘dexterity’ handling that slowed down time completion without impairing human performance. Participants in the paper-based condition, may have found easier to use pen and paper to draw the solutions, despite their impaired performance. On a behavioral level, it seems that interactive computerized systems such as CVRPS have the potential to boost human performance in optimization problem solving but also appears to slow down the time to complete CVRPs.

Also, the presence of the computerized tool may offer a task environment where participants can off-load problem complexity demands onto the tool, leading to a paradox as they have to allocate less effort to complete the problem-solving task, but at the same time put more attention to the task (promoting optimality and increasing time completion). However, to be able to further ascertain such a notion, further research is needed on this direction. In relation to the general performance across problems, the results from this experiment match with the overall prior research (e.g. Kefalidou & Ormerod (2014) and Experiment 1 in this paper as participants performed worse in the same problem in all those experiments (P4 - 39-6) even though the experimental factor was different (verbalization and cognitive load – Kefalidou & Ormerod (2014) , post-problem feedback and error-trapping – Experiment 1 and computerized interactive system use and concurrent live feedback – Experiment 2 in this paper).

5. General Discussion

There are two main threads of research contributions that the present paper offers: one on the body of literature that investigates the role and value of metacognitive support provision has on human performance in computerized problem solving and another one on the body of literature that looks upon interactive optimization. The research presented here suggests that the benefit metacognitive structures such as feedback provision offer to human performance varies depending on the context (e.g. type of task environment), on whether they co-exist with other metacognitive structures (e.g. planning and feedback provision) and on the timing of provision. Schraw (1998) suggested three main self-regulatory constructs that constitute metacognitive processes: planning, monitoring and evaluation. As discussed in the introduction and literature review in this paper, self-regulation can have a positive impact on human performance (e.g. feedback provision and planning activities – Corbett & Anderson, 2001; Cheshire et al., 2005; Schooler & Anderson (1990); Kapa (2001); Azevedo & Bernard, 1997; Kulhavy & Stock, 1989; Livingston, 1997; Shute, 2008). Results in this paper support the notion that metacognitive support mechanisms can improve human performance compared to situations where metacognitive support is not present in optimization problem solving tasks. This can be taken to further suggest that the beneficial outcomes that metacognitive support exhibits in learning and other problem solving contexts seems to extrapolate to a different domain in problem solving, that of optimization problem solving.

Error trapping, feedback provision and planning support provision (as metacognitive support) were explicitly tested within the reported experiments. The present research found that error-trapping and post-problem feedback had no effect on human performance and time completion contrary to prior research suggesting the opposite (e.g. Cheshire et al., 2005; Schooler & Anderson, 1990). However, the use of the interactive computerized support tool (CVRPS) (that facilitated metacognitive support through concurrent feedback and planning provision) improved human performance compared to performance on the pen and paper version of CVRPs where no support was present. However, no differences in human performance were found in participants' solutions that used either version of CVRPS suggesting that different metacognitive support (and self-regulatory mechanisms particularly) exhibit benefits in human performance in different ways.

More specifically, the findings from Experiment 1 suggest that as the number of nodes and routes to design was increasing (problems P4 (39-6) and P5 (45-7)), the control Group performed better than the post-problem feedback Group. Delayed feedback effects on human performance improvement were not found, and therefore, not to be able to support prior research (e.g. Schooler & Anderson, 1990), that suggests that delayed feedback improves human problem solving performance by promoting error-tracking behaviour (Delay Retention effect is not supported and Interference Preservation Hypothesis is not supported either in optimization problem solving such as CVRPs).

However, when solvers from the post-problem feedback group faced higher complexity problems, their performance degraded, yet maintained validity, providing partial support to the notion that delayed feedback can improve error-tracking. However, in problems with lower complexity, the same participants performed better than the participants from the control Group. An explanation for these results may occur due to motivational aspects. For example, error-trapping could have acted beneficially in the less complex problems (and enhancing solvers' motivations) as the provided guidance did not involve longer re-planning or 'undoing' actions leading to closer to optimal solutions. On the other hand, error-trapping for more complex problems may had impaired solvers' motivation levels (and consequently their performance) as it 'forced' them to re-think and re-plan more extensively their solutions to provide a valid solution (thus satisficing).

In such a case, higher cognitive load could be induced in more complex problems via the number of steps for re-planning and 'undoing' solvers would need to perform would also be increased. Prior research in the literature suggested that cognitive load can reduce problem solvers' motivation to act upon the provided feedback (Ashford, 1986; Corno & Snow, 1986) and as such, results from this experiment could be taken to confirm prior literature.

Similarly, in Experiment 2 the participants with a lower problem complexity performed better than those with a higher problem complexity. Despite the variability observed across the problem performances, it was not significant to suggest that error trapping and post-problem feedback had an effect on human performance. Experiment 2 results support the notion that interactive computer systems help in improving performance and solving optimization problems such as CVRPs. Results from both versions of CVRPS ('strategy/concurrent live feedback-support' and 'no strategy/with concurrent live feedback-support') show that performance in these conditions was improved when compared with performance in the paper-based condition. Similarly to Experiment 1, a significant effect of problem was found but also a significant interaction between group and completion time. This suggests that participants who used CVRPS (both ‘strategy/concurrent live feedback support’ and ‘no strategy/with concurrent live feedback support’) performed faster than those in the paper-based group. This was expected as the CVRPS tool implemented provided (automatic) support for the calculations needed. The latter – in paper-based settings- are expected usually to delay the completion of a CVRP. In relation to performance quality (how close to optimal solutions are) both CVRPS versions performed better than paper-based group. This supports the notion that interactive computerized tools can provide a layer for performance improvement (Waters, 1984; Klau et al, 2010).

Pair-wise comparisons of CVRPS performance revealed no difference between the two CVRPS versions i.e. groups showed no difference in the quality of solutions generated using the system suggesting that the task environment in these two versions did not significantly vary to generate differences in performance. Furthermore, this can be taken to suggest that the mere addition of strategy support (in the full version of CVRPS) may not have had the major influence in human performance as much as the concurrent live feedback perhaps had (which was present in both versions of CVRPS). Even though we cannot rule out the possibility that participants in this ‘no-strategy/with concurrent live feedback support’ CVRPS version were not using strategies, the tool itself did not encourage this in this condition. The fact that no significant results were found between the performance in the two CVRPS versions might suggest that the strategy support may not be critical for the quality of performance. On the contrary, the fact that both software versions generated better solutions while providing automatic feedback which is always visible on the screen (thus without necessitating to cross-check for feedback at a different place or even perform the checks themselves) might play an important role in the generation of good solutions.

More specifically, planning (strategy) support mechanisms appeared to make no difference on the quality of performance exhibited in CVRPS, contrary to computerized concurrent explanatory feedback that was present in both versions of CVRPS. This suggests that planning (strategy) support engages metacognition on a different level from that of feedback provision, which is something that can be further supported by the fact that in the CVRPS group that facilitated planning, time completion appeared to be faster compared to the CVRPS group that facilitated concurrent live feedback only. However, interesting observations did not only take place in regards to planning (strategy) provision but also in regards to the type of computerized concurrent feedback provision provided in both CVRPS groups.

Particularly, in both the stripped and full CVRPS versions, concurrent live explanatory feedback was provided but concurrent live directive feedback was only provided in the full CVRPS version alongside with the planning (strategy) support. Results show a clear benefit in human performance from the computerized concurrent live explanatory feedback, further suggesting that people attempting to find optimal solutions acquired a higher-level of problem solving capability through it. This is something that does not seem to be confirmed for the computerized concurrent live directive feedback (present in the full CVRPS version) suggesting that concurrent directive feedback may not assist people in reaching deeper engagement levels with critical components of the optimization problem solving (e.g. focusing on optimizing rather than satisficing and 'good enough' solutions). Another explanation may be that the provision of multiple metacognitive constructs support acted as an interference, providing thus another effect similar to the Interference Preservation Hypothesis according to which concurrent feedback can create interference due to error flagging. However, as the solo presence of either computerized concurrent directive feedback or planning (strategy) support was not tested in the present research within the research presented in this paper, it is not possible to further validate this notion. Future research should orientate to further explore potential effects in this direction.

The findings presented in this research are in alignment with Corbett and Anderson (2001) suggestions that computerized feedback provision in problem-solving can aid performance but does not appear to support research from Lurie and Swaminathan (2009) and Schooler and Anderson (1990) that suggest that delayed feedback is more beneficial. Furthermore, Corbett and Anderson (2001) found that immediate feedback resulted to faster responses, something that the present research found as well. It is important to note, however, that Corbett and Anderson's (2001) results referred to immediate corrective feedback (automated correction of participants' responses), which is something that was not tested in the present research. The types of feedback tested in this research included paper-based post-problem feedback (delayed feedback), computerized concurrent live explanatory feedback and computerized concurrent live directive feedback, from which computerized concurrent feedback was found to be more beneficial for participants' solutions compared to the control (no feedback) group. What this tells us is two main things: firstly, that other types of computerized concurrent feedback mechanisms appear to promote better solutions (and faster); secondly, that the beneficial computerized feedback provision impact on human performance extrapolates to other domains rather than problem solving in learning literature.

The results presented here can contribute to the Assistance Dilemma theory according to which a threshold for the level and type information (e.g. feedback) that learners and problem solvers are being provided is necessary. In the case of optimization problem solving, computerized explanatory feedback provision appears to be beneficial for promoting better human performance when it occurs during the problem solving process. It suggests that participants benefited from knowing what violations they performed at the moment they performed them and this seems that influenced positively the way they solved the problems. While planning (strategy) provision and computerized concurrent directive feedback appear to improve human performance compared to the paper-based performance, they do not appear to be better than the provision of computerized concurrent explanatory feedback. This suggests that getting information in regards to length of routes and weight loaded 'as-and-when' each route constructed, alongside with the ability to plan the routes is not enough to add benefit to the overall performance compared to the benefit gained from knowing what violations occurred.

Despite the fact that the present research did not test learning per se (but instead human performance in optimization problem solving), Assistance Dilemma is a useful theory to consider in optimization problem solving as it becomes apparent (through the present research) that different types of feedback, different types of metacognitive mechanism supports (e.g. feedback and planning) and different timings in the provision of feedback seem to impact on human performance differently.

These findings not only offer a new perspective on how feedback provision can influence performance but also on what the role of feedback is in optimization. Looking at the means of solutions quality and time completion, it seems that performance is not linear. This may occur due to problem complexity characteristics or even cognitive load effects. As such problem complexity and its relationship to cognitive load should be further explored in the context of computerized feedback provision for optimization problem solving in future research. A next step for this research would be to investigate planning mechanisms and the provision of heuristics (strategies) more explicitly in an attempt to pin down further differences amongst different metacognitive support. For example, testing for effects on human performance of planning support only was a limitation of the present research and it should be pursued in future research. Similarly, another limitation of the present research was that it did not test for potential effects of post-problem computerized feedback. Despite the present research (Experiment 1) did not find a significant effect of paper-based post-problem feedback on human performance (and thus not being further pursued in Experiment 2), computerized post-problem feedback provision is another type of metacognitive support that prior research (e.g. Schooler &

Anderson, 1990) suggested beneficial outcomes on human performance. As such, testing it in isolation, would be an avenue for future research.

Furthermore, feedback appears to be a contributor to motivation improvement in learners according to past literature (see e.g. Lepper & Chabay, 1985; Narciss & Huth, 2004). According to Anderson et al. (1984) for example, problem solvers and learners can become demoralised –or even confused- by unidentified errors and as such, feedback has to be provided either immediately, delayed or by request. Proponents of physical classrooms in learning or problem solving contexts, suggest that strategies such as team-working, self-esteem building and self-directed learning are better facilitated in face-to-face physical environments (Hattie, 1992) and this might be something that can additionally boost optimization problem solving processes. More recently, Oliver and Omari (2001) explored student performance and responses to problem solving learning tasks utilising a collaborative web-based technological approach and found that collaborative reflection can be critical for the improvement of performance. Similarly, Wang and Wu (2008) conducted a study with university students that used NetPorts, an online collaborative platform, through which they were able to process peer assessments, view anonymous peer feedback and submit homework. What they found was that high self-efficacy students tended to utilize more critical thinking and elaborate strategies while self-efficacy benefitted more from receiving elaborated feedback. In their study, student performance was improved when students were informed whether their responses were correct or not.

A future avenue for further research would be to look upon motivational, self-efficacy and effort aspects in optimization problem solving and investigate the role of feedback provision. For example, future research could test explicitly for self-regulation using e.g. the CRESST model (O'Neil & Herl, 1998) to investigate the four components of planning, self-checking, self-efficacy and effort when solving optimization problems such as CVRPs.

5.1 Towards Route Optimization Joint Cognitive Systems

Several school of thoughts have emerged positing opportunities that technology-enhanced interventions can offer for enhancing human problem-solving including *cognitive technologies* (Larkin, McDermott, Simon & Simon, 1980; Pea, 1985, Sweller, 1989; Woods & Roth, 1988), *intelligent technologies* (Salomon, Perkins & Globerson, 1991; Sleeman & Brown, 1982) and cognitive tools (Lajoie & Derry, 2013; Jonassen, 1992). The core notion that these approaches encompass is the joint-cognitive approach where machines support human operations and actions within their working or physical environment. As a consequence of these approaches, computational technologies have been viewed as extension of human cognition - and further as an extension of cognitive tasks. Within this application context, humans become supervisors rather than controllers. Particularly in problem-solving, and with the use of joint cognitive systems, humans are able to utilise their knowledge and co-operate with the computer to perform the problem-solving task in a dynamic way simulating human tutoring behavior. However, often becomes evident that the existing knowledge base does not become activated and utilised 'as and when' needed. Indeed prior research has found that humans need explicit prompts to utilise prior knowledge successfully within a problem solving context and task even if the *learning experience* is recent (Gick & Holyoak, 1980; Kotovsky, Hayes & Simon, 1985).

Investigating human thinking processes when attempting to optimize could prove to be extremely useful in the development and improvement of CVRP optimization algorithms, leading to solutions that are closer to optimal. Prior research in interactive optimization (e.g. Waters, 1984; Klau et al., 2010) investigated how human operators can synergize with computational algorithms in solving optimization problems such as VRPs, however, they did not test upon effects of feedback provision while allowing for construction of routes from scratch by humans. While the present research did not incorporate the use of any computational algorithms (as the above mentioned prior research), it has demonstrated that humans are capable of generating good solutions and that interactive environments that allow for full completion of problems with support from concurrent computerized feedback act beneficially in assisting them generating better and faster solutions. The present results provide support to the notion that computer environments such as CVRPS that provide metacognitive support can aid human performance in optimization problem solving and that can potentially act as a 'learning' platform for enhancing human cognition.

Furthermore, the present research can also provide a framework for improving computational optimization algorithms while understanding better cognitive processes when attempting to solve such hard problems (including deep machine learning approaches). For example, human performance is known to be subjected to ecological parameters and the task environment (e.g. Newell & Simon, 1956). Newell and Simon (1956), task environment and ecological rationality play an important role in successfully choosing the right heuristics at the right time for the right problem. Results from Experiment 2 suggest that planning (strategy) support sped up performance but did not necessarily improve the quality of the performance. This may be taken to imply that planning (strategies), which is something employed by a number of different computational algorithms, may not be the critical construct for optimizing human performance in this type of problem solving. Instead what the present research found was that concurrent feedback provision seems to promote self-awareness, goal-awareness and situation-awareness at the right time and at the right level leading to better performance. A further investigation on the nature of heuristics and strategies employed (e.g. Calculating and Clustering) becomes mandatory to ascertain further cognitive and metacognitive mechanisms within the context of optimization problem solving and feedback.

To conclude, the research presented in this paper has provided an insight into how a computerized interactive system can utilize both planning (strategies) support and concurrent live feedback to improve human performance when solving CVRP problems. Through this research it has been confirmed that people can find close to optimal solutions on CVRPs without the use of any computational algorithms but also it was found that concurrent live feedback provision (contrary to paper-based post-problem feedback – Experiment 1) can significantly improve performance, slowing down time completion rates. Interestingly, and in regards to the different feedback provision types utilized in the present research, it seems that computerized concurrent directive feedback and planning support do not improve human performance further compared to when participants received computerized concurrent explanatory feedback only.

Implications for future optimization interactive tools and even decision-support tools across different domains should take into account the following: 1) provision of concurrent feedback that can either warn or prevent (in case of critical systems) humans from error-making – in such an instance, the interactive system becomes the supervisor and the human is the actor utilizing the strengths that humans can offer to the optimization problem-solving process. Switching the roles in interactive systems that support a 'human-in-the-loop' approach changes the size of the problem space, an action that has the potential to promote optimality 2) flexibility in role and switching between different metacognitive support (feedback vs planning (strategy) support) could be beneficial especially when applications refer to more than one different problem-solving (or decision-making) contexts. A future step would be to include a set of additional human-inspired strategies and heuristics to test further whether planning (strategy) support within an interactive system improves performance significantly. Such an investigation would give further light as to what the significance of concurrent live feedback and planning (strategy) support provision is in improving human performance. Another future step would be to incorporate a computational algorithm routine (e.g. Tabu search) and investigate how human solvers interact with it (having available as well human-inspired strategies as in this presented research). Finally, a future step would be to investigate collaborative interactions and problem-solving as opposed to solo problem solving using the CVRPS presented here. Such research would allow us to ascertain human behaviors especially in regards to feedback provision and interpretation of useful prompts and warnings. This research has shown that the use of interactive software tools (again without the use of any computational algorithms) can improve human performance simply by providing concurrent computerized feedback.

Acknowledgements

This research was conducted as part of the author's PhD thesis at Lancaster University. This work was supported by EPSRC under Grant <EP/P50256X/1>. Many thanks to the Human Factors Research Group, The University of Nottingham for providing the resources for the completion of this paper. Also, many thanks are due to my PhD supervisors Prof. Tom Ormerod and Prof. Richard Eglese and to Dr. Alexander Coles for his comments on earlier versions of this manuscript.

6. References

Ashford, S. J. (1986). Feedback-seeking in individual adaptation: A resource perspective. *Academy of Management journal*, *29*(3), 465-487.

Azevedo, R., & Bernard, R. M. (1995). A meta-analysis of the effects of feedback in computer-based instruction. *Journal of Educational Computing Research*, *13*(2), 111-127.

Azevedo, R., & Hadwin, A. F. (2005). Scaffolding self-regulated learning and metacognition–Implications for the design of computer-based scaffolds. *Instructional Science*, *33*(5), 367-379.

Bangert-Drowns, R.L., Kulik, C.C., Kulik, J.A. & Morgan, M. (1991). The instructional effect of feedback in test-like events. Review of Educational Research, 61,213-238.

Blasum, U. & Hochstättler, W. (2000). "Application of the Branch and Cut Method to the Vehicle Routing Problem". Technical Report zaik2000-386, Centre of Applied Computer Science, Zentrum fur Angewandte Informatik Koln Technical Report.University of Cologne, Germany, 2000-386

Bodhani, A. (2012). Shops offer the e-tail experience, *Engineering and Technology*, Vol. 7 (5), 46-49

Bodhani, A. (July, 2012), Smart Transport, *Engineering & Technology*,. 70-73.

Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.

Burke, E. K. & Kendall, G. (2005). “Search Methodologies – Introductory Tutorials in Optimisation and Decision Support Techniques”, Springer, USA.

Cadenas, J. M., & Verdegay, J. L. (2000). Using ranking functions in multiobjective fuzzy linear programming. *Fuzzy sets and systems*, *111*(1), 47-53.

Cheng, C. D., & Kosorukoff, A. (2004, June). Interactive one-max problem allows to compare the performance of interactive and human-based genetic algorithms. In *Genetic and Evolutionary Computation Conference* (983-993). Springer Berlin Heidelberg.

Cheshire, A., Ball, L. J. & Lewis, C. N. (2005). Self-Explanation, Feedback and the Development of Analogical Reasoning Skills: Microgenetic Evidence for a Metacognitive Processing Account. Chapter in : In B. G. Bara, L.W. Barsalou & M. Bucciarelli (Eds.), Proceedings of the Twenty-Seventh Annual Conference of the Cognitive Science Society (435-441). Mahwah, NJ: Lawrence Erlbaum Associates

Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive science*, *13*(2), 145-182.

Colgan, L., Spence, R., & Rankin, P. (1995). The cockpit metaphor. *Behaviour & Information Technology*, *14*(4), 251-263.

Corbett, A. T., & Anderson, J. R. (1989, May). Feedback timing and student control in the LISP Intelligent Tutoring System. In *Proceedings of the Fourth International Conference on AI and Education* (pp. 64-72).

Corbett, A. T. & Anderson, J. R. (2001). Locus of feedback control in computer-based tutoring: Impact on learning rate, achievement and attitudes. In Proceedings of ACM CHI'2001 Conference on Human Factors in Computing Systems, 245-252

Corno, L., & Snow, R. E. (1986). Adapting teaching to individual differences among learners. *Handbook of research on teaching*, *3*(605-629).

Cortellessa, G., Cesta, A., Oddi, A., & Policella, N. (2004). User Interaction with an Automated Solver: The Case of a Mission Planner. *PsychNology Journal*, *2*(1), 140-162.

Crainic, T. G., Gendreau, M., & Potvin, J. Y. (2009). Intelligent freight-transportation systems: Assessment and the contribution of operations research. *Transportation Research Part C: Emerging Technologies*, *17*(6), 541-557.

Danzig, G.B. & Ramser, J. M. (1959) The truck dispatching problem, Management Science, 6, 81–91

Dihoff, R. E., Brosvic, G. M., Epstein, M. L., & Cook, M. J. (2004). Provision of feedback during preparation for academic testing: Learning is enhanced by immediate but not delayed feedback. *The Psychological Record*, *54*(2), 207.

Fagerholt, K., & Lindstad, H. (2007). TurboRouter: An interactive optimisation-based decision support system for ship routing and scheduling. *Maritime Economics & Logistics*, *9*(3), 214-233.

Fedra, K., & Harmancioglu, N. B. (2005). A web-based water resources simulation and optimization system. *Proc. of CCWI*, 167-172.

Ferreira, J. D. M., do Nascimento, H. A. D., & de Albuquerque, E. S. (2006, January). Interactive optimization in cooperative environments. In *Proceedings of the 2006 Asia-Pacific Symposium on Information Visualisation-Volume 60* (pp. 117-120). Australian Computer Society, Inc..

Fisher, M. L. (1994). "Optimal Solution of Vehicle Routing Problems Using Minimum K-trees", Operations Research, Vol. 42, 626-642.

Fukunaga, A. S., & Stechert, A. D. (1997, June). An evolutionary optimization system for spacecraft design. In *Industrial and Engineering Applications of Artificial Intelligence and Expert Systems: Proceedings of the Tenth International Conference* (p. 1). CRC Press.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive psychology*, *12*(3), 306-355.

Gleicher, M., & Witkin, A. (1994). Drawing with constraints. *The Visual Computer*, *11*(1), 39-51.

Goodrich, M. A., Stirling, W. C., & Boer, E. R. (2000). Satisficing revisited. *Minds and Machines*, *10*(1), 79-109.

Hattie, J. (1992). Measuring the effects of schooling. *Australian Journal of education*, *36*(1), 5-13.

Hu, T. C. (1982). Combinatorial Algorithms. Addison-Wesley.

Jonassen, D. H. (1992). Evaluating constructivistic learning. *Constructivism and the technology of instruction: A conversation*, 137-148.

Jones, C.V. (1996). Visualization and Optimization. Boston: Kluwer Academic Publishers. Boston.

Kapa, E. (2001). A metacognitive support during the process of problem solving in a computerized environment. *Educational Studies in Mathematics*, *47*(3), 317-336.

Kefalidou, G., & Ormerod, T. C. (2014). The Fast and the Not-So-Frugal: Human Heuristics for Optimization Problem Solving, In Proceedings of the Cognitive Science Conference, Quebec, Canada.

Klau, G. W., Lesh, N., Marks, J., & Mitzenmacher, M. (2010). Human-guided search. *Journal of Heuristics*, *16*(3), 289-310.

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review*, *19*(3), 239-264.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive psychology*, *17*(2), 248-294.

Kulhavy, R. W., & Anderson, R. C. (1972). Delay-retention effect with multiple-choice tests. *Journal of Educational Psychology*, *63*(5), 505.

Kulik, J.A. & Kulik, C.C. (1988). Timing of feedback and verbal learning. Review of Educational Research, 58, 79-97.

Lajoie, S. P., & Derry, S. J. (Eds.). (2013). *Computers as cognitive tools*. Routledge.

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*(4450), 1335-1342.

Lepper, M. R., & Chabay, R. W. (1985). Intrinsic motivation and instruction: Conflicting views on the role of motivational processes in computer-based education. *Educational Psychologist*, *20*(4), 217-230.

Lepper, M.R., Woolverton, M., Mume, D.L. & Gurtner, J. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In S. Lajoie & S. Derry (eds.) Computers a cognitive tools Hillsdale, NJ: Erlbaum

Lee, E. K. & Mitchell, J. E. (1998). "Branch-and-bound methods for integer programming". In the Encyclopedia of Optimization, Kluwer Academic Publishers.

Livingston, J. A. (1997). Metacognition: An Overview, http://*www.gse.buffalo.edu/fas/shuell/CEP564/Metaeog.htm*.

Lurie, N. H. & Swaminathan, J. M. (2009), “Is Timely Information Always Better? The Effect of Feedback Frequency on Decision Making,” Organizational Behavior and Human Decision Processes, Vol. 108 (March), 315-329

MacGregor, J. N., & Ormerod, T. (1996). Human performance on the traveling salesman problem. *Attention, Perception, & Psychophysics*, *58*(4), 527-539.

MacGregor, J. N., & Chu, Y. (2011). Human performance on the traveling salesman and related problems: A review. *The Journal of Problem Solving*, *3*(2), 2.

Maffioli, F. (2003). Subgradient optimization methods in integer programming with an application to a radiation therapy problem. (Unpublished Dissertation). Kaiserlautern University, Kaiserlautern, Germany.

Maringer, D. G. (2006). *Portfolio management with heuristic optimization* (Vol. 8). Springer Science & Business Media.

Michalewicz, Z. & Fogel, D. (2000), How to Solve it: Modern Heuristics, Springer Verlag.

Mu, Q., Fu, Z., Lysgaard, J., & Eglese, R. (2011). Disruption management of the vehicle routing problem with vehicle breakdown. *Journal of the Operational Research Society*, *62*(4), 742-749.

Munro, A. Fehling, M.R. & Towne, D.M. (1985). Instruction intrusiveness in dynamic simulation training. Journal of Computer-Based Instruction, 12, 50-53.

Narciss, S., & Huth, K. (2004). How to design informative tutoring feedback for multimedia learning. *Instructional design for multimedia learning*, 181-195.

Newell, A., & Simon, H. (1956). The logic theory machine--A complex information processing system. *IRE Transactions on Information Theory*, *2*(3), 61-79.

Newell, A. and Simon, H. A. (1972). Human Problem Solving. Prentice-Hall, Englewood Cliffs, NJ

Oliver, R. & Omari. A. (2001). Exploring Student Responses to Collaborating and Learning in a Web-Based Environment. Journal of Computer Assisted Learning, 17(1), 34-47

Pea, R. D. (1985). Beyond amplification: Using the computer to reorganize mental functioning. *Educational psychologist*, *20*(4), 167-182.

Poli, R., & Cagnoni, S. (1997). Genetic programming with user-driven selection: Experiments on the evolution of algorithms for image enhancement. *Genetic Programming*, 269-277.

Salomon, G., Perkins, D. N., & Globerson, T. (1991). Partners in cognition: Extending human intelligence with intelligent technologies. *Educational researcher*, *20*(3), 2-9.

Schmidt, R.A., Young, D.E., Swinnen, S. & Shapiro, D.C. (1989). Summary knowledge results for skill acquisition: Support for the guidance hypothesis. Journal of Experimental Psychology: Learning, Memory and Cognition, 15, 352-359

Schooler, L. J. & Anderson, J. R. (1990). The disruptive potential of immediate feedback. In Proceedings of the Twelfth Annual Conference of the Cognitive Science Society, 702-708, Cambridge, MA

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional science*, *26*(1-2), 113-125.

Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, *78*(1), 153-189.

Scott, S. D., Lesh, N., & Klau, G. W. (2002, April). Investigating human-computer optimization. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 155-162). ACM.

Sims, K. (1991). *Artificial evolution for computer graphics* (Vol. 25, No. 4, pp. 319-328). ACM.

Sleeman, D., & Brown, J. S. (1982). *Intelligent tutoring systems* (pp. 345-pages). London: Academic Press.

Sreevalsan-Nair, J., Verhoeven, M., Woodruff, D. L., Hotz, I., & Hamann, B. (2007). Human-guided enhancement of a stochastic local search: Visualization and adjustment of 3d pheromone. In *Engineering Stochastic Local Search Algorithms. Designing, Implementing and Analyzing Effective Heuristics* (pp. 182-186). Springer Berlin Heidelberg.

Surber, J. R., & Anderson, R. C. (1975). Delay-retention effect in natural classroom settings. *Journal of Educational Psychology*, *67*(2), 170.

Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of educational psychology*, *81*(4), 457.

Todd, S., & Latham, W. (1999). The mutation and growth of art by computers. *Evolutionary design by computers*, *9*, 221-250.

Toth, P. & Vigo, D. (2001). "Branch-and-bound algorithms for the capacitated VRP". In: Toth, P., Vigo, D. (Eds.), The Vehicle Routing Problem. SIAM: Philadelphia, 29-52.

Toth, P. & Vigo, D., (2002). The Vehicle Routing Problem, SIAM Society for Industrial and Applied Mathematics, Toth, P. and Vigo, D. (eds), Philadelphia USA

Van Vliet A., Boender C., G., E. & Rinnooy Kan A., H., G. (1992). Interactive optimization of bulk sugar deliveries. *Interfaces*, 22 (3), 4–14

Waters, C.D.J. (1984). Interactive vehicle routeing. *J. of the Operational Research Society*, 35(9), 821-826

Woods, D. D., & Roth, E. M. (1988). Cognitive engineering: Human problem solving with tools. *Human Factors*, *30*(4), 415-430.