

Augmenting On-Screen Instructions with Micro-Projected Guides: When it Works, and When it Fails

Stephanie Rosenthal,^{1,3} Shaun K. Kane,^{2,3} Jacob O. Wobbrock,² Daniel Avrahami³

¹School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213
srosenth@cs.cmu.edu

²The Information School, DUB Group
University of Washington
Seattle, WA 98195
{skane, wobbrock}@u.washington.edu

³Intel Labs Seattle
1100 NE 45th Street
Seattle, WA 98105
daniel.avrahami@intel.com

ABSTRACT

We present a study that evaluates the effectiveness of augmenting on-screen instructions with micro-projection for manual task guidance unlike prior work, which replaced screen instructions with alternative modalities (*e.g.*, head-mounted displays). In our study, 30 participants completed 10 trials each of 11 manual tasks chosen to represent a set of common task-components (*e.g.*, cutting, folding) found in many everyday activities such as crafts, cooking, and hobby electronics. Fifteen participants received only on-screen instructions, and 15 received both on-screen and micro-projected instructions. In contrast to prior work, which focused only on whole tasks, our study examines the benefit of augmenting common task instructions. The augmented instructions improved participants' performance overall; however, we show that in certain cases when projected guides and physical objects visually interfered, projected elements caused increased errors. Our results demonstrate that examining effectiveness at an instruction level is both useful and necessary, and provide insight into the design of systems that help users perform everyday tasks.

Author Keywords

Task guidance, computer-assisted instruction, augmented reality, micro-projection, computer vision, everyday tasks.

ACM Classification Keywords

K.3.1 [Computing Millieux]: Computer assisted instruction, H.5.1 [Information interfaces and presentation]: Artificial, augmented, and virtual realities.

ACM General Terms

Human Factors, Experimentation

INTRODUCTION

The use of instructions for completing everyday manual

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

UbiComp '10, Sep 26–Sep 29, 2010, Copenhagen, Denmark.
Copyright 2010 ACM 978-1-60558-843-8/10/09...\$10.00.

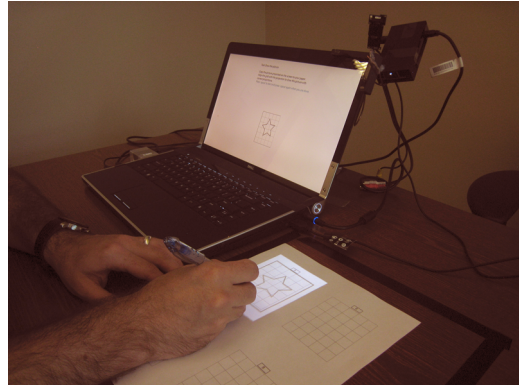


Figure 1. Augmenting on-screen instructions with micro-projection for guiding a tracing activity.

tasks is an integral part of life. We follow recipes when cooking, use maps when driving, consult manuals for putting furniture together, or learn the rules of a board game. On the web, a growing number of sites are dedicated to user-generated instructions, guides, recipes, how-tos, and do-it-yourself projects. As the availability of these resources grows, people increasingly turn to their computers for instructions to help them complete various manual tasks in the real world. Still, interpreting, following, and applying instructions can often be a challenge.

Prior work in the field of Augmented Reality (AR) has shown that instructions superimposed onto the physical space through a head-mounted display (HMD), are effective for task guidance (*cf.* [3, 5]). Other work used large, top-mounted projectors with additional instrumentation to guide users through physical tasks [8]. Recent advancements in micro-projection technology have made it possible to imbue typical laptop computers with small projection capabilities (*cf.* [10]). While micro-projectors are limited in many respects compared to HMDs (*e.g.*, in projection area, resolution, and brightness), we propose that using micro-projection to augment, rather than replace, on-screen instructions for task guidance can deliver some strengths of in-place guides with the benefits of high-resolution screen-based instructions, all in a portable and affordable way.

In this paper, we present a study that examines in detail the effectiveness of augmenting on-screen instructions with a micro-projector pointed at the table for computer-based task

guidance (Figure 1). We compared users' performance in a variety of common tasks (involving activities such as assembling, sorting, organizing, and folding) as users followed instructions presented on a laptop screen or a combination of on-screen instructions and micro-projected guides. Our results show that people were able to complete tasks 27% faster *overall* and with 32% fewer errors when using the combined instructions compared to using screen-only instructions. Our results also show, however, that projected instructions can *harm* performance, specifically when they interfere visually with physical objects.

Our work provides three primary contributions: First, our study examines the effectiveness of augmented instructions at the *task-component level* (separately considering components required for a variety of manual tasks such as locating items, positioning items, identifying errors, *etc.*), and in doing so provides a detailed understanding of how and when augmented instructions should be used within the process of an entire task. Second, to our knowledge, ours is the first study to show that in certain cases projection is not only unhelpful, but can *harm* performance. Third, our work extends prior work on see-through head-mounted displays by showing the usefulness of systems that combine low-fidelity in-place micro-projection and an upright high-resolution screen for providing task guidance.

RELATED WORK

Our investigation of the effectiveness of augmenting on-screen instructions with micro-projection for task guidance draws from research in the areas of Computer-Aided Instruction, remote instructions in Computer Supported Cooperative Work (CSCW), and Augmented Reality.

Computer-Aided Instruction: Years of research in the area of Computer-Aided Instruction (CAI) have shown that computer-based instructions can support effective learning. CAI findings suggest adapting instructions to a user's current level of comprehension as well as using video, animation, and other media to improve information retention [1, 6, 18]. While multimedia demonstrations have been shown to increase task performance when compared to static text or verbal instructions, studies have found that demonstration information is not retained for long [16, 23]. To address this limitation, Palmiter and Elkerton [16] suggest providing both text instructions and "well-placed" instructions to give the immediate benefits of multimedia while supporting retention. Our work builds on this notion by providing both screen-based instructions along with "well-placed" guides projected upon the task area. Our current study focuses on static text and pictorial instructions on a screen and micro-projector, although we expect that integrating multimedia instructions would only bring added benefits that have been described in prior CAI literature.

Remote Instruction in CSCW: The use of projection for guiding users through tasks is not entirely new. A line of research in CSCW examines different mechanisms to allow two or more people who are not co-located to perform tasks

together. Particular attention is given to cases where one person remotely guides another person through a task (for example, a medical expert remotely guiding a local doctor through a complex medical procedure). As previous studies have shown [8, 13, 17], technology that enables a helper to see into the remote workplace can improve coordination, allow participants to maintain a shared context of the task, and use shared language to describe the task effectively. Other research [7, 11, 12] has introduced interactive capabilities to the remote viewing experience, allowing the helper to point, highlight, or gesture into the remote space. Fussell *et al.* [7] found that superimposing a remote helper's sketches over a user's work area can significantly reduce the time needed to complete a task. Our work draws upon this prior research and aims to provide similar improvements to the use of computer-guided instructions.

Task Guidance in Augmented Reality: Finally, and most directly, our work builds on and extends prior research on the use of Augmented Reality (AR) for guiding users through manual tasks. Most existing work on task guidance in AR uses wearable high-resolution head-mounted displays (HMDs) for presenting information in space. Feiner *et al.*'s [5] early work demonstrates the use of AR for guiding a user through a complex laser printer repair by overlaying visual instructions on the view of the printer. Recent work [14, 15, 19] demonstrates AR for a variety of tasks, which include playing musical instruments, reading, and manufacturing. Ju *et al.* [8] presented the Origami Desk, an interaction design exhibit that uses top-mounted projection, electric field sensing, and swept-frequency sensors to instruct users in folding origami paper into boxes and cranes. (While inspiring, it provides only anecdotal evidence as to the effectiveness of its approach.)

Studies of the appropriateness of HMD-based instructions for task guidance over on-screen instructions have found AR to be effective. Neumann and Majoros [15] identify several cognitive benefits that reduce error rates when using AR for task guidance. For example, overlaying instructions onto the physical objects reduces the need to memorize instructions before performing the task. Baird and Barfield [2] show that participants who used a see-through HMD in a circuit board building task were at least 47% faster at completing the task than when they used paper or on-screen instructions. Tang *et al.* [20] compared the effectiveness of spatially-aligned AR instructions, delivered via an HMD, to paper and on-screen instruction in a LEGO construction task. Their study showed a significant benefit of spatially-aligned instructions, primarily over paper instructions.

Unlike with see-through HMDs, for which display area is effectively infinite and full instructions can be displayed, and unlike for top-mounted projectors (such as those used by Ju *et al.* [8]) that have large projection area and can produce high-fidelity images, micro-projection imposes limitations on both projection area and resolution. We therefore refer to our projected material as "guides." One implication of these differences between HMDs, top-

mounted projections, and micro-projections is that the benefits of HMDs for task guidance, as shown in prior work, may not directly apply to task guidance using other tools such as micro-projection.

Prior work [2, 5, 20] compared the effectiveness of AR instructions at the level of an entire task. However, we cannot assume that all components of a task (e.g., locating the next piece vs. assembling that piece in the right manner) would benefit equally from AR instructions. A taxonomy of assembly instructions or task-components, proposed by Dunston and Shin [4], describes how AR could help in categories such as “*Identify*” a specific piece and “*Cut materials*.” While originally created for designing AR instructions in the area of civil engineering, we use this taxonomy to investigate a hybrid system for presenting instructions for everyday physical tasks, such as cooking, electronics, and arts and crafts.

This review of prior work raises a number of interesting questions: Could the advantages shown for the use of projectors with remote instructions in CSCW be useful for CAI? Would the ability to use both an upright high-resolution screen and low-fidelity micro-projected guides prove useful? Could a simple small laptop-mounted micro-projector deliver today some of the promise of future head-mounted AR systems? Which task-components or categories of instructions benefit most from augmentation? Our study answers some of these questions, fills gaps in our understanding of the role of augmented instructions in physical-task guidance, and informs the design of systems that provide guidance for everyday manual tasks.

STUDY DESIGN

To investigate the effectiveness of a system that combines a laptop with micro-projection for task guidance, we designed a study that compares screen-based instructions augmented with micro-projected guides (*Screen+Projection*) to screen-only instructions (*Screen-Only*) typical in laptop use. To understand the benefits and limitations of the two modes of instruction, we designed instructional material for a variety of manual task-components (11 overall), drawing from 7 of the 9 categories proposed by Dunston and Shin [4] (the last two categories relate to clarifying communication and are not relevant here). We should note that a projection-only condition was not included because the limitations of micro-projectors in projection-size and resolution mean that such a condition would be guaranteed to be inferior to both the *Screen-Only* and *Screen+Projector* conditions (it would be impossible to fit both the textual instructions as well as guides in the projection). Including a projection-only condition with low-resolution technology would therefore not have provided meaningful insights about the performance benefits of projected guides.

The tasks and categories used in this experiment, shown in Table 1, were: *Identify Piece*: Locate a physical item in a set based on its shape and size. *Identify Missing Piece*: Given a set of 10 Scrabble tiles and a corresponding list of

11 letters on the screen/projector, identify the missing letter. *Trace Shape*: Using pen and paper, draw a specified shape to scale. *Position Wire*: Insert a piece of electric wire into a hole in a breadboard specified by a diagram. *Mark Location*: Using pen and graph paper, mark the cell specified by a corresponding diagram. *Cut Wire*: Use scissors to cut a piece of 8” wire into three segments of specified length. *Fold Paper*: Fold a piece of origami paper once based on a diagram. *Bend Wire*: Bend an 8” wire to match a simple rectilinear shape. *Sculpt*: Roll out 2 oz. of Play Doh into a tube of a specified length, or spread it to fit a specified area. *Verify Breadboard*: Given a breadboard and a diagram, verify that the location of a wire already in the breadboard matches the diagram. *Verify Grid*: Compare a printed 5×5 grid of symbols to a displayed grid and verify that the two grids match.

Our experimental tasks closely mirror everyday manual tasks. For example, our *Identify Piece* task echoes instructions to find the correct size screw when assembling furniture. For each of the tasks, we created 10 trials of approximately equal difficulty, as well as a single, simplified practice trial. For each trial, we designed both on-screen instructions (seen by all participants), and micro-projected guides (seen by half of the participants).

METHOD

In order to compare the two modes of instruction, we conducted a between-subjects experiment. Each participant performed 1 practice and 10 real trials for each task described above, for a total of 110 real trials per subject. We compared the error rates and completion times between display conditions for each task.

Apparatus

On-screen instructions were presented on a Dell Studio XPS 16 laptop running Windows Vista. An Optoma DLP micro-projector running at a resolution of 480×320 was mounted to the top-right of the laptop screen using a plastic bracket. A Logitech webcam was directed at the table to the right of the laptop and was connected to the experimenter’s computer to view the participants’ work. We used black tape to mark a working area on the table that was fully within the camera view. Participants were asked to perform the tasks within the working area when possible. Figure 1 shows the experimental setup with micro-projection.

For each trial, we created static images for the on-screen and projected instructions. We created custom software using Python and OpenGL to both display the images and record when a key was pressed to start or end a trial.

Procedure

Participants were assigned at random to one of the two display conditions (*Screen-Only* or *Screen+Projection*). The order in which tasks were presented to a participant was determined using a Latin Square. We ensured that each row in the Latin Square was completed by two participants, one from each display condition.


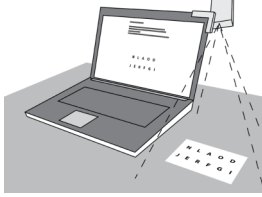

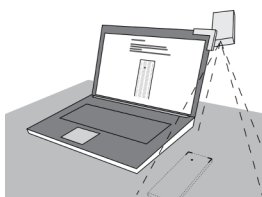
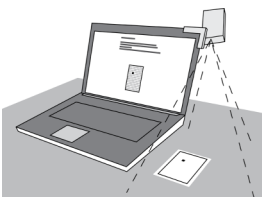

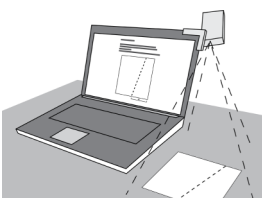
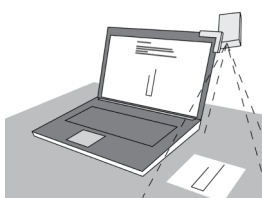
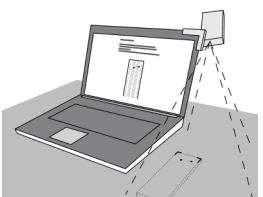
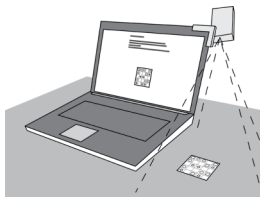
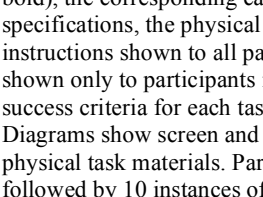
<p>Identify Piece</p> 	<p><i>Identify in Set</i></p> <p>Task: Identify a plastic piece from a bin of objects by size and shape. Materials: Bin containing 10 plastic circles, squares, or trapezoids. On Screen: Outlines of all shapes in a bin with target shape highlighted. Projected: Outline of target shape, to scale. Success: Correct piece identified. Real-World Example: Identify the correct screw to build a desk; identify correct resistor for a circuit.</p>	<p>Identify Missing Piece</p> 	<p><i>Identify in Set</i></p> <p>Task: Identify a missing piece from a set of tiles. Materials: Ten Scrabble game tiles in bin, each with a single letter on it. On Screen: The 10 letters from the tiles, plus 1 additional letter, in a random order. Projected: The same as the screen. Success: Missing tile identified. Real-World Example: Confirm all puzzle pieces are in the bin; determine if pieces missing from assembly set.</p>
<p>Trace Shape</p> 	<p><i>Draw/Trace</i></p> <p>Task: Trace specified shape accurately onto grid paper. Materials: Pen and a paper with 5x6, 1/2" cell grids. On Screen: A shape, drawn over a 5x6 grid (e.g. heart, star, spiral). Projected: The same as the screen, to scale. Success: Trace remains within $\pm 1/4$" of the specified shape. Real-World Example: Teach kids how to write letters or draw cartoons; Write "Happy Birthday" on cake evenly.</p>	<p>Position Wire</p> 	<p><i>Placement (fine)</i></p> <p>Task: Insert a wire into a specified hole in breadboard. Materials: A breadboard & a piece of wire. On Screen: A diagram of a breadboard with 1 highlighted hole. Projected: Alignment markers for board, a circle around the correct hole. Success: Wire placed in the correct hole. Real-World Example: Copy complex pattern or design; Assemble circuits on breadboard or PCB.</p>
<p>Mark Location</p> 	<p><i>Placement (coarse)</i></p> <p>Task: Draw an X in specified cell on a sheet of grid paper. Materials: A piece of paper with 10x15, 1/4" cell grids and pen. On Screen: A 10x15 grid drawing with one cell highlighted. Projected: The same as the screen, to scale. Success: The correct cell was marked. Real-World Example: Determine where screws go in furniture; Interactively teach chess or checkers indicating possible moves.</p>	<p>Cut Wire</p> 	<p><i>Cut</i></p> <p>Task: Cut piece of 8" wire into 3 segments of specified length. Materials: 8" wire, scissors, and ruler. On Screen: 3 lines, with the segment length specified in inches. Projected: 3 lines, to scale. Success: Each segment of wire is within $\pm 1/4$" of the specified length. Real-World Example: Cut yarn for art project; Mark length of bread or rolls to divide the dough evenly for baking.</p>
<p>Fold Paper</p> 	<p><i>Reshape</i></p> <p>Task: Fold a sheet of paper along specified line. Materials: 1 piece of 15cm x 15cm Origami paper. On Screen: A square with a dotted line for the fold, and numbers showing the width of the fold (e.g. 1/3 the paper length). Projected: Same as screen, to scale. Success: Both ends of the fold are within $\pm 1/4$" of the specified line. Real-World Example: Teach origami indicating folds.</p>	<p>Bend Wire</p> 	<p><i>Reshape</i></p> <p>Task: Bend 8" wire to match a specific shape. Materials: An 8" wire for each trial. On Screen: A line representing the wire, with up to 2 right-angle bends. Length of sides not given explicitly. Projected: Same as screen, to scale. Success: Wire has the correct number of segments, each within $\pm 1/4$" of specified length. Real-World Example: Indicate the curve of model airplane wings.</p>
<p>Verify Breadboard</p> 	<p><i>Find Errors (fine)</i></p> <p>Task: Verify that wire in breadboard with matches on-screen diagram. Materials: A breadboard with wire, pen, and grading sheet. On Screen: Diagram of board with wire drawn and 2 holes highlighted. Projected: Board corner markers, circles over two wire holes. Success: Correctly identified where physical wire-ends don't match diagram. Real-World Example: Check for errors in complex circuits.</p>	<p>Verify Grid</p> 	<p><i>Find Errors (coarse)</i></p> <p>Task: Identify extra, conflicting, missing symbols on paper grid compared to on-screen grid. Materials: Pen, and a 5x5 paper grid containing 15 symbols: circles, squares, and stars. On Screen: 5x5 grid with up-to 2 differences from paper. Projected: Same as screen. Success: Correctly marked cells where the paper and screen/projected differ. Real-World Example: Check furniture assembly.</p>
<p>Sculpt</p> 	<p><i>Spread/Cover</i></p> <p>Task: Roll and spread Play Doh modeling dough into a tube of a specified length, or sheet of a specified length and width. Materials: 2 oz of Play Doh, ruler. On Screen: A line or rectangle drawing, with lengths in inches. Projected: Same line or rectangle, to scale. Success: Sculpted shape is within $\pm 1/4$" of the specified dimensions. Real-World Example: Roll cookie dough to size.</p>		

Table 1. The 11 tasks used in the study. Includes the task name (in bold), the corresponding category name (from [4]), task specifications, the physical materials provided, the on-screen instructions shown to all participants, the projected instructions shown only to participants in the *Screen+Projection* condition, the success criteria for each task, and real-world examples of the task. Diagrams show screen and projected instructions, but not the physical task materials. Participants completed one practice trial, followed by 10 instances of a task for each of the 11 tasks. Each participant thus completed a total of 110 trials.

Participants were told that time and accuracy would be measured for each task and that they should strive to complete each task as quickly and accurately as possible. For each task, the participant was first asked to read the instructions and ask any questions they had. They were then given a practice trial for the task. The completion time and accuracy of the practice trial were not recorded. Once ready, the participant started each of the 10 trials by pressing the spacebar, and ended each the same way. Each participant completed all 10 trials before moving to the next task. The order of trials within a task was fixed for all participants. After finishing all 11 tasks, participants completed a questionnaire, and were paid \$20 for their time. The study took approximately one hour.

Participants

Thirty participants were recruited for the study using Craigslist and university mailing lists for staff and students. There were 17 female participants (10 in the *Screen-Only* condition). Nine participants were left-handed (four in the *Screen-Only* condition). The average age was 27.4 (min 20, max 46, $SD=7.2$) Nineteen participants reported they were students. Other occupations included technical support staff, office worker, homemaker, librarian, and document manager. Participants performed a total of 3,300 trials, 1,650 in each display condition.

MEASURES

We logged key presses and collected task artifacts for all trials. From these, we computed two performance measures: time and error rate. To compute a task's error rate, each trial was first scored as either a success or an error using the criteria described in Table 1. A wide variation in how accurately participants performed the *Bend Wire* task made scoring by a single person difficult. For this task only, we used multiple coders as follows: After scoring 30 trials as a group, three of the authors each scored 2/3 of the remaining 270 trials independently, such that each trial was scored twice. This method resulted in a high Kappa agreement of $\kappa = 0.85$. Conflicting scores (29 of 270) were resolved by a third coding.

The following variables were used in the analysis:

Dependent Variables:

- **Time (log₁₀-transformed):** The time to complete a trial, in milliseconds. Since completion time displayed a log-normal distribution, it was transformed prior to analysis.
- **Error Rate:** The percentage of error trials (out of a possible 10) in each task.

Independent Variables:

- **Display Condition:** *Screen-Only* vs. *Screen+Projection*.
- **Task-Type:** The 11 tasks described in Table 1.
- **TrialID:** The name of an individual trial.
- **TaskOrder:** The order (1-11) in which a task was presented to the participant.
- **TrialOrder:** The order (1-10) in which a trial within each task was presented to the participant.

Covariates:

- **ParticipantID**
- **Age**
- **Gender:** *Female* vs. *Male*.
- **Handedness:** *Left* vs. *Right*.

ANALYSIS

We utilized two separate statistical models to analyze the effects of display condition and task type on Time and Error Rate.

Completion-time was analyzed using a mixed-effects model analysis of variance in which Time (log₁₀-transformed) was the dependent measure. Display Condition (*Screen-Only* vs. *Screen+Projection*), Task-Type, TaskOrder (1-11), TrialOrder (1-10), Handedness (*L* vs. *R*), and all two-way interactions were modeled as fixed effects. ParticipantID and TrialID were modeled as random effects.¹

Error-rate was analyzed at the task level using Poisson regression for event-count data [22] in which task-level Error Rate (percent in task) was the dependent measure. Display Condition (*Screen-Only* vs. *Screen+Projection*), Task-Type, TaskOrder (1-11), Handedness (*L* vs. *R*), and all two-way interactions between them were modeled as fixed effects.² Finally, to control for individual differences, ParticipantID was nested within Handedness and Display Condition, and included in the model as a fixed effect.

Since neither Age nor Gender showed a significant relationship with Time or Error Rate, we did not include them in the models.

RESULTS

Our results show that adding micro-projected guides can lower completion time and error rate overall. However, a detailed look revealed that not all tasks benefited from the added projected guides and in one case, performance was, in fact, *worse* for participants using the guides. Both the completion-time and error-rate models were significant ($p < .001$). The model for completion time had a high R^2 of 0.77, meaning that the model was able to account for 77% of the variance in the data.

Display Condition and Task-Type

Display Condition had a significant main effect on both Time ($F[1,26]=12.9$, $p < .002$) and Error Rate ($\chi^2(1)=27.6$, $p < .001$) with faster completion time for *Screen+Projection*

¹Trial ID was modeled as a random effect because we are not interested in the specific 10 trials we generated per task, but in the larger set of possible trials that these were drawn from, just as we are interested in the population from which participants were drawn, not the specific individuals who participated.

²The infrequency of errors resulted in a highly non-normal distribution precluding the option of using an ANOVA. Poisson regression is typically utilized to model rare-event data like this.

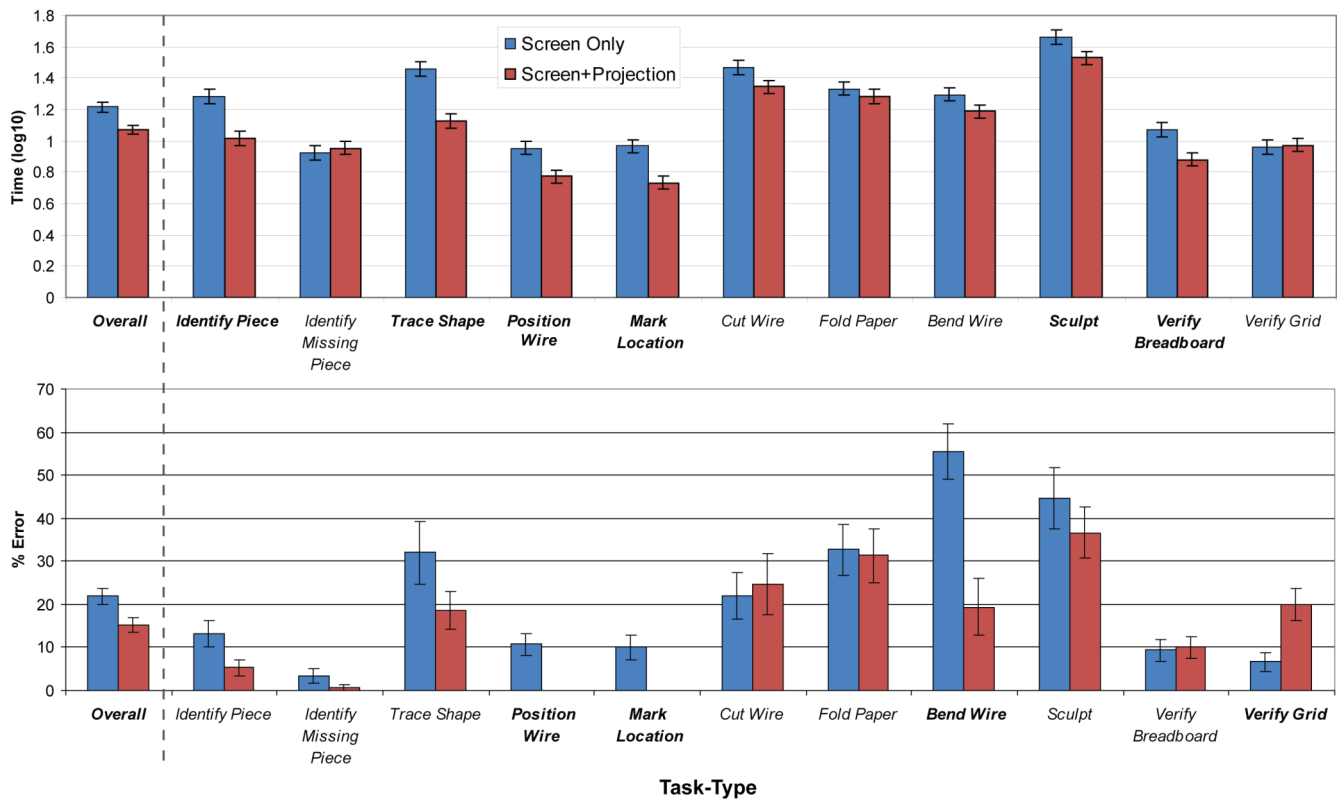


Figure 2. The main effect of Display Condition on completion time (log₁₀ scale) and error-rate and effect by Task-Type. Lower values are better. Tasks shown in bold indicate a statistically significant difference between display conditions.

(12s vs. 16.5s)³ as well as fewer errors per task (15% vs. 22%). These main effects replicate results from prior work in AR [2, 20] that show the benefit of augmented instructions over entire tasks (Figure 2).

As one would expect due to the qualitatively different nature of the various task-components, Task-Type also had a significant main effect on both Time ($F[10,92]=58.1$, $p<.001$) and Error Rate ($\chi^2(10)=263.1$, $p<.001$). For example, the *Sculpt* task took 39 seconds on average to complete, while a *Mark Location* averaged only 7 seconds.

Our analysis now departs from prior studies by examining differences between conditions over individual task-components. Interactions between Display Condition and Task-Type were also significant for both Time ($F[10,3125]=35.4$, $p<.001$) and Error Rate ($\chi^2(10)=65.3$, $p<.001$), implying that the way in which micro-projected guides affected completion time and error rate was not equal for all task types. In order to identify tasks in which combined instructions differed significantly from screen-only instructions, we examined completion time and error rate for each Task-Type, comparing between *Screen-Only* and *Screen+Projector*. We first conducted a planned pair-

wise comparison of completion Time between display conditions for each task type. To account for the number of pair-wise comparisons (11) and reduce the risk of Type I statistical errors, we employed a Bonferroni correction, resulting in $\alpha=0.0045$, *i.e.*, a difference is considered statistically significant if and only if $p<.0045$.

As can be seen in Figure 2, in 6 of the 11 tasks, participants in the *Screen+Projector* condition completed each trial significantly faster, on average, than participants in the *Screen-Only* condition. Those tasks were *Verify Breadboard* (8s vs. 12s; $F[1,37]=19.7$, $p<.001$), *Identify Piece* (10s vs. 19s; $F[1,37]=36.6$, $p<.001$), *Sculpt* (34s vs. 46s; $F[1,37]=9.7$, $p<.004$), *Trace Shape* (13s vs. 29s; $F[1,37]=58.2$, $p<.001$), *Position Wire* (6s vs. 9s; $F[1,37]=17.3$, $p<.001$), and *Mark Location* (5s vs. 9s; $F[1,37]=27.9$, $p<.001$). These results support the effectiveness of augmenting on-screen instructions with micro-projected guides, especially because these 6 tasks represent 5 of the 7 task-categories from [4] that we evaluated. No other comparisons were significant.

Next we conducted similar pair-wise comparisons for Error Rate (again, with $\alpha=0.0045$). Four task types showed significant difference in Error Rate between the display conditions (see Figure 2, bottom). We found that in both of the tasks from the *Placement* task category (*Mark Location* and *Position Wire*), participants in the *Screen+Projector* condition not only completed trials significantly faster than

³Because the independent variables were not completely orthogonal, we used Least Squared Means (LS Means) to control for the values of the other independent variables. When possible, the means reported are LS Means.

participants in the *Screen-Only* condition, but that all participants completed these tasks without any errors. This is compared to an average error rate of 10% for *Mark Location* for participants in the *Screen-Only* condition ($\chi^2(1)=18.5$, $p<.001$) and 11% for *Position Wire* ($\chi^2(1)=18.4$, $p<.001$). Considering the precision demands of the *Position Wire* task, the result is somewhat surprising and supports the use of micro-projected guides for similar *Placement* tasks. Error Rate in the *Bend Wire* task was much higher for participants in the *Screen-Only* condition (55%) than participants in the *Screen+Projector* condition (19%; $\chi^2(1)=21.6$, $p<.001$).

Finally, a surprising result shows that participants in the *Screen+Projection* condition did significantly worse in the *Verify Grid* task than participants in the *Screen-Only* condition (6.7% vs. 20.0%; $\chi^2(1)=8.9$, $p=.003$). Recall that in this task, users were asked to compare a printed 5×5 grid of symbols to a displayed (and projected) grid, and identify missing, extra, or conflicting symbols on the paper.

In order to understand the possible causes for this negative effect (we are unaware of any prior reports of a negative impact of augmented instructions), we analyzed the 300 *Verify Grid* trials in greater detail. We first coded the number and type of mismatches (*Symbol Missing from Paper*, *Extra Symbol on Paper*, and *Conflicting Symbol on Paper*) in each trial. We then analyzed the interaction between Display Condition and Type of Mismatch on the likelihood of error in each trial to determine whether the type of mismatch was associated with different error levels for the different display conditions. As the interaction was significant ($F[2,236]=4.9$, $p<.01$) we then performed a Tukey HSD pair-wise comparison (see Figure 3).

The results show that when a symbol on paper conflicted with a displayed/projected symbol (*Conflicting Symbol on Paper*), participants in the *Screen+Projection* condition committed significantly more errors than participants in the *Screen-Only* condition ($F[1,185]=18.6$, $p<.001$). There was no significant difference between display conditions when a symbol was *Missing from Paper* ($F[1,85]=0.9$, $n.s.$), nor when there was an *Extra Symbol on Paper* ($F[1,122]=3.3$, $n.s.$). This indicates that *Screen+Projection* participants' errors can be attributed to significantly worse performance at identifying conflicting symbols (to illustrate, consider a contrived example where the letter "F" is projected on top of the letter "E"). This finding indicates that the spatial interaction between the physical and virtual must be considered. We conclude our results with a description of the effects of control measures and covariates.

Handedness

There was no significant effect of Handedness on completion time ($F[1,26]=0.2$, $n.s.$). However, left-handed participants had a significantly higher error rate than right-handed participants (23% vs. 16%; $\chi^2(1)=11.0$, $p<.001$). Looking closely, we found that the average error rate was actually marginally higher for left-handed participants in

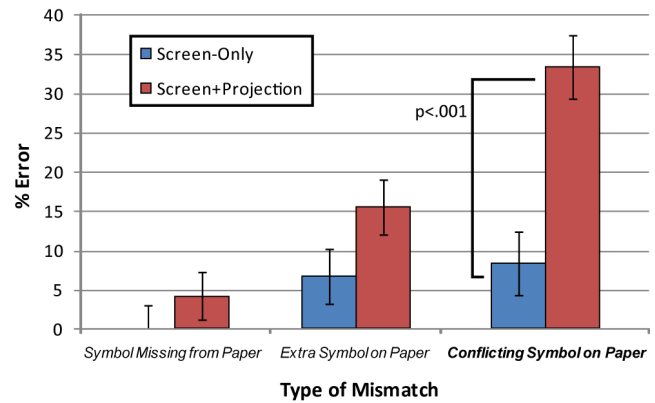


Figure 3. Errors in *Verify Grid* by Display Condition and Type of Mismatch. Significant difference for *Conflicting Symbol*.

the *Screen-Only* condition (25%) than in the *Screen+Projection* condition (22%; $\chi^2(1)=3.7$, $p=.054$). We can thus conclude that the increased error rate for left-handed participants was not due to occlusion of the projection, and more likely due to the requirement to work on the table immediately to the right of the laptop (possibly inconvenient for left-handed users).

Learning Effects

Somewhat surprisingly, we found neither a significant main effect of TrialOrder on Time ($F[1,91]=0.7$, $n.s.$), nor any significant two-way interactions with TrialOrder. TaskOrder did have a significant effect on Time ($F[1,3125]=10.6$, $p<.002$), with participants completing tasks faster as the study progressed. The interaction between Task-Type and TaskOrder was significant for Time ($F[10,3140]=7.2$, $p<.001$), indicating that tasks encountered at different stages of the study resulted in different improvements in completion time. It is interesting to notice that the shorter completion time did not translate to a significant increase in Error Rate ($\chi^2(1)=0.2$, $n.s.$), discounting the possibility that participants became careless over time. No other effects or interactions were significant.

To summarize, our results show significant overall benefit, in completion time and error rate, of the *Screen+Projection* condition. A closer examination shows that this benefit was not true in all cases, and in one case (*Verify Grid*) the projection harmed performance. An additional analysis revealed that this was likely a result of visual interference between projected and physical (printed) elements.

DISCUSSION

Like prior work in AR, the micro-projected guides improved participants' overall performance. A closer examination showed that in 7 of the 11 tasks, performance with the guides lowered completion time, reduced error rate, or both. In 3 tasks, performance did not significantly differ, and in 1 task, performance was, in fact, worse. By considering the categories of tasks, we can understand which ones are aided by augmented instructions and should be used within the process of entire tasks.

In tasks for which projected guides improved performance, we observed benefits similar to those previously reported [2, 20], including reducing visual context switching by providing to-scale feedback, and by reducing the need for additional tools. These benefits of micro-projected guides are encouraging considering the micro-projector limitations discussed above, and were made possible through the additional availability of on-screen instructions. We next discuss these benefits, as well as drawbacks.

Benefits of Combined Instructions

As we cannot assume that augmented instructions would benefit all task-components equally, we hypothesized that projected guides would improve performance on some task-components more than others. For task-components that require the user to check many details from the instructions, like debugging a circuit or drawing a complicated pattern, the user may have to glance back and forth between the screen and work area many times. By placing feedback directly on the work area, the projected guides removed the need to switch their gaze between the screen and table. For example, in the *Identify Missing Letter* task, participants were required to look at the screen to see the full set of letters, then down at the table to confirm which letter tiles were present. In contrast, we observed many participants in the *Screen+Projection* condition set their tiles adjacent to the projection area so that they could view the letter list and the tiles simultaneously. Although this effect did not result in significantly faster performance in the *Identify Missing Letter* task, it likely contributed to the performance of this and other tasks, such as *Trace Shape* and *Position Wire*, which are used often in art and engineering projects.

In real-life tasks such as model airplane or furniture building, it is important to identify the correct pieces to use. The projected guides gave participants the ability to match physical objects to the to-scale projections. In many cases, this allowed the participant to replace a difficult mental operation with a much simpler one. For example, in the *Identify Piece* task, *Screen-Only* participants often removed all of the pieces from a bin and placed them on the table, while participants in the *Screen+Projector* condition could simply match single pieces to the size of the projected shape (Figures 4a and 4b). In our experimental design, we chose to use on-screen images that were not to scale, as instructional images found on the Web are typically not to scale. Thus, presenting images to scale on screen might improve performance for some tasks, although this would clearly not be possible for tasks where the user needs to place something in front of the screen (e.g., *Mark Location*), or use a vertical surface (e.g., *Sculpt*).

Finally, an additional benefit of to-scale projection was the ability to reduce the need for additional tools, specifically the ruler. In cake decorating, for example, projecting a ruler could help space “Happy Birthday” evenly when it is not feasible to place an actual ruler in the icing. All participants were given a ruler during those tasks that required precise

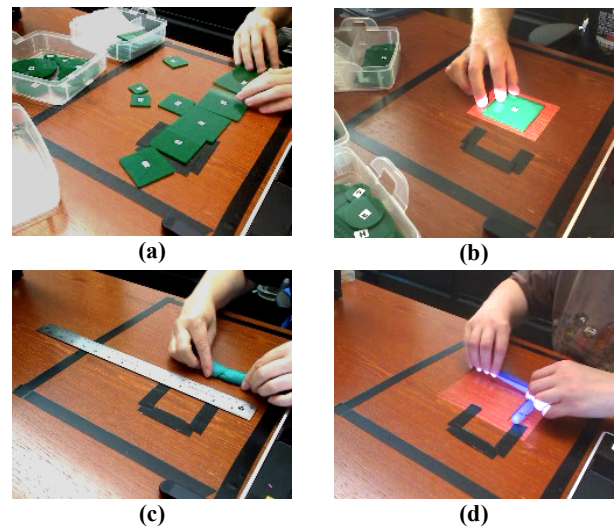


Figure 4. Participants performing the *Identify Piece* task (top) and *Sculpt* task (bottom). (a) A participant in the *Screen-Only* condition arranges all of the pieces on the table to compare them. (b) A participant in the *Screen+Projection* condition matches a piece to its projected outline. (c) A participant in the *Screen-Only* condition uses a ruler to measure Play Doh. (d) A participant in the *Screen+Projection* condition uses a projected guide to measure length while sculpting.

measurement (*Cut Wire*, *Sculpt*). Participants in the *Screen-Only* condition were required to perform an additional measurement step with the ruler, while participants in the *Screen+Projector* condition were able to use the to-scale projected measurements, removing the need to switch tools and perform an additional step (Figures 4c and 4d). Although our tasks used only the ruler as a measuring tool, to-scale projection would likely reduce the need for other measuring tools, such as a compass or grid paper.

The combined screen and projection system was also more flexible for the participants’ different work styles and provided two sources of information to complete the tasks. We found that some participants seemed to use the screen primarily to verify the projected information. For example, it is not feasible to project all of the holes in the breadboard but the screen instructions clearly show every hole. After using projected guides to insert a wire into the breadboard, some participants checked their work by comparing it to the on-screen breadboard diagram. Although the on-screen and projected information in this condition were similar, future systems could use the screen to present other information that cannot be displayed in the projection area, such as high-resolution images or multi-media demonstrations.

Because the combined system was implemented on a laptop, as opposed to a static table setup, it supports users’ workflow by allowing them to perform tasks where they prefer. A user could bring it to a workshop to assemble electronics or other materials and into the kitchen for help with cooking. While we tested the system in only a single environment, there was nothing that would have prevented us from moving the setup to different places for each task.

Limitations of Micro-Projected Guides

Although participants performed faster and more accurately with the projected guides for some tasks, many participants encountered difficulties using the projections. In the *Verify Grid* task, in particular, participants had difficulty distinguishing projected material from the printed material we provided. We also observed the following problems with use of the projected guides: visual interference, hand occlusion, a limited workspace, and difficult aligning real-world objects to the guides. These problems are common to both high-resolution and low-fidelity projection systems.

In the *Verify Grid* task, the projected shapes were overlaid directly onto the printed shapes for participants to verify whether the two sets of shapes were identical. It was difficult for participants to see the output of the projector on top of the printed grid. Participants in both the *Screen+Projection* and *Screen-Only* conditions had similar difficulty identifying when there was an extra shape on the printout that was not in the instructions. Participants in the *Screen+Projection* condition were significantly worse at detecting when two overlapping shapes were not the same. Some participants adapted to this by placing the paper at a slight offset to the projection to compare the shapes, but this issue likely contributed to the increased error rate for that task. The difficulty separating the augmented information from the underlying object may affect all augmented reality systems [2], but could perhaps be addressed by projecting in an area adjacent to the grid, or projecting in colors that stand out better against the object.

Some participants also experienced difficulties when their hand occluded the projected guides—a common problem in front-projection systems. This happened most often when participants moved their hands into the projection area to work on a task, such as *Trace Shape*. While participants could often resolve the occlusions themselves, one potential automated solution is to detect possible occlusions, and redirect the projected information to another area of the table, possibly using a callout as in *Shift* [21].

Another common problem was the limited workspace provided by the projector mounted on a laptop monitor. Even with a small area, we see a lot of benefit in the combined screen and projected instructions. However, the workspace size required participants to adopt some specific poses to complete tasks and use the projected guides. For example, during the *Cut Wire* task, participants in the *Screen-Only* condition typically held the wire in front of themselves to cut it, while participants in the projector condition often held the wire down to the table to cut it to match the projected guide. The location of the workspace was also difficult for left-handed participants, who had to reach across the laptop to the right side to complete the tasks. These problems could be addressed by providing a steerable projector, by providing projected feedback even if the user's hand is outside the projection area, or by mounting two projectors on the laptop, one on each side as in *Bonfire* [10].

Finally, some participants in the *Screen+Projection* condition experienced some difficulties when aligning a physical object to the projected guides. This was not a significant problem in the *Trace Shape* or *Fold Paper* tasks, where the projection clearly showed the outline of the object, but was challenging in the breadboard tasks, where a slight misalignment of the breadboard could result in inaccurate projected guidance. As discussed previously, a benefit of the combined screen and projection system is that participants could use the screen instructions to verify their alignment. This problem should be addressed in future systems by using a camera to detect the object's position and orientation, and adjusting the projection to match.

Task Completion Strategies

Our robust system enabled participants to use a variety of (sometimes unexpected) strategies for completing the tasks. For example, in the *Identify Piece* task, participants in the *Screen-Only* condition used various strategies to quickly find the correct shape. Some participants laid out their shapes in the same pattern as shown on screen, while others stacked the shapes by decreasing size, and some picked up shapes a few at a time to make pair-wise comparisons. In another example, some participants used the ruler to create straight edges in the *Sculpt* task by cutting the Play-Doh using the ruler's edge. One major design implication of this result is that task guidance systems should allow users to perform tasks in a variety of ways without disrupting the overall flow of instruction. Although we initially expected that participants in the *Screen+Projection* condition would line up the letter tiles with the projection in the *Identify Missing Letter* task, very few in fact did so.

FUTURE WORK

The current study provides evidence of the effectiveness and potential pitfalls of augmenting on-screen instructions with micro-projected guides for a number of categories of task-components. However, future work should examine the appropriateness of this feedback for more complex, multi-step tasks that utilize these components. In some cases, the tasks here can be extended to these multi-step tasks. For example, our *Fold Paper* task could be combined to teach a user to create an origami (similar to Ju et al. [8]). Supporting multi-step tasks will require additional interaction methods to step through the task, for example, animation portraying or giving feedback for a dynamic process, and could include an authoring tool for recording new instructions. Further analysis is needed to determine which components of multi-step tasks are helped most and which might be harmed by augmented instructions.

Another item for future work is using computer vision to detect and recognize the actions a user is taking, and possibly adjust the guides to correct missteps or change task orderings based on user behaviors. A simple example would be a vision-based inventory of small parts, whereby after the user lays out all the parts necessary for a given task, the system informs the user of any missing pieces.

Finally, future research in this area should also explore the use of different types of on-screen and projected feedback. The instructions evaluated in this study featured text and static images only on a high-resolution laptop and low-fidelity projection. Including animation, video, interactive components, and higher fidelity projections could potentially improve both on-screen and projected guides.

CONCLUSION

We have presented a user study that compared task performance using typical on-screen instructions to a system that augments on-screen instructions with micro-projected guides. The results of the study show that combined instructions can improve speed and reduce errors for a variety of task-components present in manual tasks. However, we also found that projected guides resulted in worse performance when projected guides and physical objects visually interfered with each other. These results extend prior research on AR-based instruction by showing that task guidance can be implemented effectively using commodity hardware available today (a laptop, webcam, and micro-projector). Furthermore, by evaluating users as they perform a large set of task-components, we have shown where projected guides can provide benefits within a variety of tasks that could be augmented in the future.

ACKNOWLEDGMENTS

The authors thank Matthai Philipose, Anind Dey, Adam Rea, Darren Gergle, Yaakov Kareev, Pedja Klasnja, Anthony LaMarca, Richard Beckwith, and Beverly Harrison.

REFERENCES

1. Anderson, J. R., Corbett, A. T., Koedinger, K. R., Pelletier, R., Cognitive tutors: lessons learned. *Journal of the Learning Sciences*, 4, 2 (1995), 167-207.
2. Azuma, R., Bailiot, Y., Behringer, R., Feiner, S., Julier, S., Macintyre, B. Recent advances in augmented reality. *Computer Graphics and Applications*, IEEE Press, 21, 6(2001), 34-47.
3. Baird, K.M., Barfield, W., Evaluating the effectiveness of augmented reality displays for a manual assembly task. *Virtual Reality*, 4, 4 (1999), 250-259.
4. Dunston, P.S., Shin, D. H., Key areas and issues for augmented reality applications on construction sites. In *Mixed Reality in Architecture, Design, and Construction 2009*, Springer (2009). 157-170.
5. Feiner, S., Macintyre, B., Seligmann, D., Knowledge-based augmented reality. *Communications of the ACM*, 36, 7 (1993), 53-62.
6. Fletcher-Flinn, C. and Gravatt, B. The efficacy of computer assisted instruction (CAI): a meta-analysis. *Journal of Educational Computing Research*, 12, 3 (1995), 219-242.
7. Fussell, S.R., Setlock, L. D., Yang, J., Ou, J., Mauer, E., Kramer, A. D. I., Gestures over video streams to support remote collaboration on physical tasks. *Human-Computer Interaction*, 19, 3 (2004), 273-309.
8. Gergle, D., Kraut, R.E., and Fussell, S.R. (2004). Action as language in a shared visual space. In *Proc. CSCW 2004*, ACM Press (2004), 487-496.
9. Ju, W., Bonanni, L., Fletcher, R., Hurwitz, R., Judd, T., Post, R., Reynolds, M., and Yoon, J., Origami Desk: integrating technological innovation and human-centric design. In *Proc DIS 2002*, ACM Press (2002), 399-405.
10. Kane, S.K., Avrahami, D., Wobbrock, J. O., Harrison, B., Rea, A., Phillipose, M., LaMarca, A., Bonfire: a nomadic system for hybrid laptop-tabletop interaction. In *Proc. UIST 2009*, ACM Press (2009), 129-138.
11. Kirk, D., Fraser, D. S., Comparing remote gesture technologies for supporting collaborative physical tasks. In *Proc. CHI 2006*, ACM Press (2006), 1191-1200.
12. Kirk, D., Rodden, T., Fraser, D. S., Turn it this way: grounding collaborative action with remote gesture. In *Proc. CHI 2007*, ACM Press (2007), 1039-1048.
13. Kraut, R. E., Miller, M. D. & Siegel, J. (1996). Collaboration in performance of physical tasks: effects on outcomes and communication. In *Proc. CSCW 1996*, ACM Press (1996), 57-66.
14. Motokawa, Y., Saito, H., Support system for guitar playing using augmented reality display. In *Proc. ISMAR 2006*, IEEE Press (2006), 243-244.
15. Neumann, U., Majoros, A., Cognitive, performance, and systems issues for augmented reality applications in manufacturing and maintenance. In *Proc. Virtual Reality*, IEEE Press (1998), 4-11.
16. Palmiter, S., Elkerton, J., An evaluation of animated demonstrations for learning computer-based tasks. In *Proc. CHI 1991*, ACM Press (1991), 257-263.
17. Ranjan, A., Birnholtz, J. P., Balakrishnan, R., An exploratory analysis of partner action and camera control in a video-mediated collaborative task. In *Proc. CSCW 2006*, ACM Press (2006), 403-412.
18. Swets, J.A., Feurzeig, W., Computer-aided instruction. *Science*, 150, 3696 (1965), 572-576.
19. Taketa, N., Hayashi, K., Kato, H., Noshida, S., Virtual pop-up book based on augmented reality. In *Interacting in Information Environments*, Springer, 2007, 475-484.
20. Tang, A., Owen, C., Biocca, F., Mou, W., Comparative effectiveness of augmented reality in object assembly. In *Proc. CHI 2003*, ACM Press (2003), 73-80.
21. Vogel, D. and Baudisch, P. 2007. Shift: a technique for operating pen-based interfaces using touch. In *Proc. CHI 2007*, ACM Press (2007), 657-666.
22. Vermunt, J.K. *Log-linear Models for Event Histories*. Thousand Oaks, California: Sage Publications, 1997.
23. Watson, G., Curran, R., Butterfield, J., Craig, C., The effect of using animated work instructions over text and static graphics when performing a small scale engineering assembly. In *Proc. CE 2008*, Springer (2008), 541-550.