# How's Your Sewing? Investigating Metrics to Automatically Assess Sewing Expertise

Marcel Lahaye
lahaye@cs.rwth-aachen.de
RWTH Aachen University
Aachen, Germany

Ricarda Rahm
rahm@cs.rwth-aachen.de
RWTH Aachen University
Aachen, Germany

Andreas Dymek
andreas.dymek@rwth-aachen.de
RWTH Aachen University
Aachen, Germany

Adrian Wagner
wagner@cs.rwth-aachen.de
RWTH Aachen University
Aachen, Germany

Judith Ernstberger
judith.ernstberger@rwth-aachen.de
RWTH Aachen University
Aachen, Germany

Jan Borchers
borchers@cs.rwth-aachen.de
RWTH Aachen University
Aachen, Germany

## ABSTRACT

Makers must regularly assess their expertise when planning projects or selecting tutorials. However, personal bias makes this assessment prone to error, potentially leading to frustration, loss of materials, and discouragement. Additionally, hobbyists have limited feedback possibilities to refine their skills, unlike, for example, apprentice artisans who receive continuous instructor feedback. To address these issues, automated expertise assessment systems could help makers assess their skills and progress. However, such systems require assessment metrics, which have been studied little in the maker context so far. We derived such metrics for sewing from semi-structured interviews with ten sewing-related instructors about their evaluation process. Additionally, we showed them a sewn object and asked them to assess the creator's expertise. From our findings, we derive criteria to use in future automated sewing expertise assessment systems. For one criterion, seam allowance, we present a functional demonstrator that automatically assesses related measurements.

## CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

## KEYWORDS

Expertise Assessment, Sewing

**ACM Reference Format:**
Marcel Lahaye, Ricarda Rahm, Andreas Dymek, Adrian Wagner, Judith Ernstberger, and Jan Borchers. 2024. How's Your Sewing? Investigating Metrics to Automatically Assess Sewing Expertise. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3613905.3651067

Permission to make digital or hard copies of part or all 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. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
*CHI EA '24, May 11–16, 2024, Honolulu, HI, USA*
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0331-7/24/05
https://doi.org/10.1145/3613905.3651067

## 1 INTRODUCTION

When planning a crafting project, makers often need to assess their expertise in the skills required to complete that project [1, 33]. In sewing, for example, we learned from our interviewees that velvet is harder for novices to handle than cotton due to its slipperiness. Therefore, sewers need to determine whether they can use this material successfully or need to substitute it with an alternative.

Beyond just completing projects, makers also strive to learn and improve their skills [18]. Frequent feedback possibilities can enhance learning [34, 35] and potentially increase motivation by, for example, demonstrating continuous improvement [11]. While, for instance, artisan apprentices can consult their instructors about project planning and feedback, hobbyists usually have limited access to such resources. Local or online communities can provide alternative support but vary widely in availability and competencies [31]. Further, domain experts who are not trained in tutoring tend to misjudge the expertise required for tasks [14] and how to phrase expertise-appropriate statements [15]. Overall, this can lead to project execution issues, potentially resulting in adverse effects [29] like cognitive overload [22, 25], frustration [21], and material waste. While some makers consider this a learning opportunity, others can even be discouraged from pursuing their craft [33].

To address these issues, we envision support through automated expertise assessment systems. Such systems could provide more at-hand crafting expertise feedback, similar to research on automated expertise assessment in, e.g., software [2, 9, 13], games [5], sports [3, 19], and surgery [17, 26, 27]. Such automated assessment requires domain-specific metrics that model user expertise, like a climber's ascent speed modeling their climbing experience [19].

To initiate and support research in this direction, we identified potential metrics for both activity- and result-based assessment in sewing. Our metrics are based on interviews with professional sewing-related instructors who regularly need to evaluate the skills of their students. We analyzed those interviews and identified multiple parameters we expect to be usable by an automated assessment approach. Finally, we chose one such parameter, *seam allowance*[1], and implemented a functional software prototype that assesses this metric similarly to established sewing teaching practice.

[1] "Seam allowance refers to the area between the stitching and raw, cut edge of the fabric" https://www.thesprucecrafts.com/seam-allowance-definition-2978260, (accessed Mar 22, 2024)

With our initial contribution, we intend to instigate our and others' automated crafting expertise assessment research to ultimately support makers in assessing and improving their skills easily and autonomously.

## 2 RELATED WORK

The terms *expertise*, *skill*, and *competence* are used interchangeably in HCI research, with similar definitions for all [5, 9]. Expertise is distinguished into *declarative knowledge* and *procedural knowledge*. While the first are facts and rules that can be written down and communicated more easily, the latter refers to knowledge about how to do something, which is mostly learned through practice [13, 24]. Apparently, *procedural knowledge* about a skill differentiates the novice from the expert [24]. Expertise can further be modeled as a network of weighted nodes in which the weight represents the probability of the user possessing the knowledge represented by that node [8].

Our ultimate goal is to support makers in assessing and improving their skills autonomously, enabling them to plan and complete their projects successfully. Therefore, in our case, we refer to crafting expertise as the ability to successfully complete a crafting project to one's own expectations. This considers different notions of "success" between makers from just wanting to "make it work" to "perfecting" their craft [32]. Therefore, instead of presenting expertise optima, crafting expertise assessment systems could emphasize the relative progression of the user's skill improvements. This would enable user agency in deciding when their individual target expertise is reached.

Automated expertise assessment systems assess the user's expertise based on the activity or its result, using various, mostly domain-specific metrics that indicate an expertise level in the related domain [17]. Activity-based assessment considers metrics during the activity, like the steadiness of paintbrush strokes. Result-based assessment utilizes quality metrics of the result of an activity, like the evenness of a surface coat of paint.

Metrics are usually based on the software or object that the user interacts with or controls, the tools used during the activity, or body measurements. Examples of such metrics are software-specific *command efficiency* [13], software-independent *number of opened submenus* [9, 16], or tool-specific *maximum mouse velocity and acceleration* [2, 9] and *average number of keys pressed at one time* [5]. Alternatively, systems can quiz the user to assess their expertise [6, 8].

Automated sport expertise assessment often uses activity- and body-related metrics, such as the activity-related *swimming velocity* [3], body-related *climbing stability* [19], and tool-related *golf club swing movement* [12].

Surgery expertise assessment uses metrics like forces and torques applied to tools [26, 27] or accelerometer data from tool movement during surgery [17].

Otherwise, while most automated assessments rely on domain-specific metrics, a domain-independent approach can be attempted using a generalized assessment algorithm based on activity accelerometer data and expert assessments [17]. This would not require domain knowledge at the system design stage.

In the crafting domain, Gong et al. present an analysis that combines data from multiple sensors placed around a workshop, on tools, and on the user. One use case presented is a classifier trained on the user's fabrication activities using ratings from two experts [10]. They report this as an initial possibility to automatically assess expertise in multiple fabrication-related tasks. Leake et al. utilize the user's self-reported expertise to augment sewing patterns with matching practice tasks [20]. However, substituting the self-reported assessment with a system-based expertise assessment could improve the expertise match of the practice tasks by reducing potential self-assessment bias [23].

Overall, we conclude that to enable automated crafting expertise assessment, measurements that model the user's crafting expertise need to be identified.

## 3 METHODOLOGY

The focus of our study was to understand how instructors in the domains of sewing and tailoring assess their students' expertise to derive concepts for an expertise assessment system. To this end, we conducted semi-structured interviews online via video call with ten sewing-related instructors (Tab. 1) and asked them about their assessment criteria and guidelines for student evaluations. Participants were recruited by contacting sewing education facilities such as Universities of Applied Sciences in Germany. Therefore, results could be biased regarding national assessment customs. However, for our goal of an initial impression of teaching practice, we assume the impact of this bias to be low.

We further asked about any grading systems or examination forms and how our participants assess student progress in general. We also enquired about the remote learning situation during the COVID-19 pandemic since we expected remote assessment to be particularly informative for automated expertise assessment due to the use of digital tools.

To also simulate potential result-based expertise assessment, we showed instructors a small bag tailored by one of the authors (Fig. 1). The author self-assesses as a novice and has tailored as a hobbyist occasionally. We asked instructors to assess the maker's expertise, tell us on what criteria they based their assessment, and categorize the maker roughly into either novice, intermediate, or expert.

We applied thematic coding [7] to identify themes and metrics in the answers and understand the sewing domain vocabulary. For an opening understanding, we used *initial coding*, and *focused coding* [30] to infer the assessment metrics. The codes, themes, and findings were shared and discussed regularly between two of the authors.

## 4 RESULTS

We identified three major themes: assessing a person's skill, assessing product quality, and sewing learning approaches. The first two contain notes about the instructors' approaches to assessing the student's skill or the quality of the product created. The last theme, 'sewing learning approaches', contains instructors' comments about potential differences between curriculum-based learning and a project-based self-taught approach. In the latter, the learner attempts projects they are interested in and learns by trial and error, a common approach in the hobbyist community [33]. Questions about instructors' opinions on this were part of our interview because


**Figure 1: Small bag, sewn by one of the authors, that was assessed by the instructors during the interviews. b) Front view, c) side view, and d) inside-out view. Instructors mainly addressed the uneven zipper attachment at the top (a) and the small *seam allowance* at the bottom (e). The *seam allowance* is the spacing between the edge of the fabric and the stitching line.**

| ID | Proficiency |
|---|---|
| 1 | Teaches introductory courses at a university |
| 2 | Professor of pattern construction |
| 3 | Professor for clothing technology |
| 4 | Research Associate teaching clothing technology |
| 5 | Head of a tailoring workshop at a UAS |
| 6 | Head of a tailoring workshop at a UAS |
| 7 | Teaches manufacturing at a UAS |
| 8 | Teaches costume conception at a School of Arts |
| 9 | Head of a tailoring workshop at a UAS |
| 10 | Conducts sewing classes at a UAS |

**Table 1: Individual proficiencies of our study participants. UAS stands for University of Applied Sciences.**

we were curious about potential evaluation differences between these approaches to deduct implications for automated assessment. Afterward, we marked individual assessment criteria as subthemes and separated them into activity- and result-based criteria, following the reporting used in related work [17]. Table 2 & 3 show the resulting filtered lists. All identified metrics are individually referenced in this result section. Interviews were conducted in German; participant quotes have been translated to English.

## 4.1 Activity-Based Sewing Expertise Assessment

Being able to use a sewing machine is a major part of sewing expertise. One instructor mentioned that one of the first indications they use to see whether a person knows how to use a sewing machine is whether they can insert the threads into the machine correctly because novices often struggle with this activity (M1.3). In P3's class, the aim is to set up a machine in three minutes, and students are timed while setting up their machine. Instructors can thus recognize beginners because they need longer for each step due to their uncertainty around the machine (M1.2):

> P1: "You can see this [expertise with the sewing machine] relatively quickly. [...] everything is approached quite slowly with a lot of attention."

Two instructors mentioned that they can even discern a user's sewing machine expertise just by listening to the sound of the machine (M1.1). Further, students must understand multiple machine settings and how to set them correctly (M1.3). For example, they need to know how to set the thread tension, what happens if it is incorrect, and which needles to use for which material (M1.4).

After setting up the machine, students are assessed on their sewing technique. Three instructors mentioned that in starting courses, students train to sew on a line by doing so on a sheet of paper with varying printed lines. Over time, these lines get more complex. Students are assessed based on how accurately they can sew on a line (M2.1) and later whether they can sew at a constant distance to it (M2.2):

> P1: "We first sew a line on a sheet of paper. There are templates from companies that you first practice on to sew onto this line."

However, one instructor advised against sewing on paper:

> P8: "Sewing on paper, I don't think much of that. That is completely nonsensical and idiotic because we do not sew costumes out of paper. For costumes, we sew garments out of fabric."

Instructors also examine students on their ability to plan ahead. One instructor asked students to provide a technical drawing of their planned, tailored object to assess the plan together. The plan is then annotated together on online whiteboard tools and first test-cut at a smaller scale on actual material (M3).

However, instructors also mentioned that students often prefer a 'learning by doing' approach:

> P1: "I think people approach it differently. We have students who say they are not going to use a pattern right now but rather drape the fabric on a mannequin and just look at it and see how it falls. Then they will pin everything together and sew it in place."

Further, six instructors mentioned aspects of sewing knowledge they consider fundamental, a good indicator of a good student, and a necessary baseline for good results and wasting less material. However, instructors also noted that students understandably appreciate early successes to stay motivated and, especially for projects for themselves, are often already satisfied if the piece merely fulfills its purpose, disregarding whether all seams are sewn straight and cleanly. Makers in the DIY community often value fast iterations and creative processes, and for some, it is enough if their project just works [18, 32], which should be considered when designing crafting expertise assessment systems.

Finally, sewing theory is tested in exams (M4). Examples include differences in material properties, which machines and needles to use for which material, and how to identify error sources. We categorized this as an activity-based assessment metric because, according to instructors, increased sewing knowledge improves sewing execution.

## 4.2 Result-Based Sewing Expertise Assessment

As part of the expertise assessment, students present a result that is evaluated by instructors. According to one instructor, students receive points for properties like outer processing, interior processing, and choice of fabric. Another metric is how cleanly material is cut. Clean cutting lines improve project quality (M5):

> P1: "If the cut is not straight, how can the piece be straight? First, it has to be cut nicely and straight, and that's how one thing builds upon another."

Similarly, instructors looked for an even *seam allowance*, meaning that the sewn line is at a constant distance from the material edge, with the spacing depending on the material (M6.1 & M6.2). All but one instructor mentioned the ability to sew a constant *seam allowance* as an indicator of expertise. If the *seam allowance* is too small or omitted, the seam can tear, or the clothing size can be incorrect. These characteristics are similar to the properties *sewing on a line* (M2.1) and *constant distance* (M2.2) of the activity-based metric *sewing technique*. However, assessment with these metrics can either be done during the sewing activity or afterward using the result. Therefore, we categorize this into both activity- and result-based assessment.

Further mentioned criteria of the seam were the *seam haptic*, *seam puckering*, *seam slippage*, *seam strength*, and whether an effort was made to hide the seam (M7.1–M7.5).

Different materials are more suitable for different expertise levels (M8). For example, smooth fabrics that move around while sewing them are hard for novices. Not only are different materials suitable for different expertise types, but the same is true for different sewing projects (M9). It was recommended that students start with simple objects like a phone case, an apron, or a bag. Clothes are more complex to tailor than bags, for example, because they need to fit and have a smaller margin of error (M10). As an additional example, the German apprenticeship guidelines for tailoring[2] suggest that students have to learn to tailor clothes in a specific order. Additionally, specific clothing attachments like a T-shirt collar or a clothing hem seem to indicate that a person has more expertise in sewing (M11).

> P10 [referring to the bag shown in the interviews]: "This isn't a beginner because a zipper was used."

Finally, the expertise of the maker of the presented bag (Fig. 1) was assessed as between novice and intermediate. This was mainly based on the small *seam allowance* (M6) at the bottom and the misaligned attachment of the zipper at the top (M5). However, the ability to attach the zipper (M11), in general, was evaluated as increased expertise.

| ID | Activity-based metric |
|---|---|
| M1 | Working with a sewing machine |
| M1.1–M1.4 | Sound, Speed, Setup, Troubleshooting |
| M2 | Sewing technique |
| M2.1 | Sewing on a line |
| M2.2 | Constant distance |
| M3 | Pattern creation |
| M4 | Theory knowledge |

**Table 2: List summarizing activity-based the metrics identified in our interviews.**

| ID | Result-based metric |
|---|---|
| M5 | Clean lines |
| M6 | Seam allowance |
| M6.1 | Even |
| M6.2 | Correct distance |
| M7 | Seam quality |
| M7.1–M7.5 | Haptic, puckering, slippage, strength, visibility |
| M8 | Material |
| M9 | Project type |
| M10 | Clothing size |
| M11 | Attachment/Augmentation type |

**Table 3: List summarizing the result-based metrics identified in our interviews.**

# 5 DISCUSSION

From our interviews, we identified lists of both activity- and result-based metrics for assessing a user's sewing expertise (Tab. 2 & 3). Result-based assessment allows supporting the user outside of the sewing activity. For example, previously used materials can indicate familiarity with them and the techniques required to process them. Similarly, successfully tailoring snugly fitting clothing can indicate advanced expertise. The absence of such completed projects in the user's history could suggest a sewing novice. In contrast, activity-based assessment metrics enable live feedback during sewing. The

[2]https://www.gesetze-im-internet.de/mschnausbv_2004/index.html (accessed Mar 22, 2024)

sewing machine, one of the main tools used in sewing, was mentioned frequently. This makes augmenting that machine with technology to track sewing activity a promising approach.

## 6 AUTOMATED SEWING ASSESSMENT PROTOTYPE

To demonstrate the applicability of our findings, we created a prototype iOS app that assesses one of our metrics, *seam allowance* (M6). This is the distance between the edge of the fabric and the stitching line. For assessment, the user prints a sewing practice worksheet from inside the app, sews onto it, trying to keep a constant distance from the given line as in existing training practice, and takes a photo of the result with the app (Fig. 2). Our app uses established algorithms from libraries and related work as described below:

We first use the built-in iOS CoreImage rectangle detector[3] to correct photo distortion. Using OpenCV 4.8.0[4], we then perform color thresholding and apply a connected components algorithm to recognize each of the two lines and reduce image noise. For this, the printed and sewn lines must use different colors, in our case red and green. Each line is then represented as an array of the coordinates of its pixel points. Next, we calculate how well the user replicated the line and how constant the distance between the two lines is. For this, we aim to find a close match overlap of both lines to identify related pairs of pixel points from each line that can be used to measure distances. Assuming that the user roughly followed the printed line, we can apply an established *iterative closest point* (ICP) algorithm [4, 28] in 2D to find an optimized overlapping alignment of both lines. We test for the smallest *chamfer distance* over the full point sets by overlapping and translating the center pixel point of the sewn line onto a range of pixel points around the center of the target line. Afterward, we compute the resulting list of pixel point neighbors with the closest *Euclidian distance*.

We use this to calculate the *replication accuracy* (M6.1 even) and *distance variation* (M6.2 correct distance). The first is the percentage of overlapping pixel points between the target and the sewn line. We use a threshold overlap distance of six pixels based on a test with an automated stitching machine to offset potential margin errors in the photo. A high *replication accuracy* indicates that the user was able to replicate the target line accurately. The *distance variation* is the standard deviation of the distances between the point pairs of the lines. A low *distance variation* indicates that the user could sew their line at a constant distance from the target line. We derived these two values as expertise measurements from our interviews. Good scores in these two measurements imply that the user is advanced in stitching a *seam allowance* (M6) at a constant distance along a pre-cut piece of fabric. However, such scores might be understood differently between makers [32]. Therefore, we suggest usage of our metrics for self-assessment and an individual's progressive improvements.

This workflow provides a low-cost training possibility to hobbyists that does not consume potentially costly pieces of fabric, one of the main deterrents from practicing sewing [20]. However, one of our interviewees also criticized training with paper since it behaves differently than fabric, potentially reducing the training effect for fabric. To address this, automated stitching machines like the Bernina 8 series[5] could be utilized to stitch the provided training pattern onto scrap material. Alternatively, the training pattern could be drawn by hand, potentially guided by a template. However, the application uses the known length of the practice worksheet lines to provide absolute evaluation values. Therefore, the user would need to attempt to draw these patterns at the same scale, or the evaluation values would need to be considered relative. Other known limitations are closed shapes like a circle: Stitching at a constant distance around a circle would result in a larger shape with little possible overlap with the target circle. While the *distance variation* could still be calculated, the resulting *replication accuracy* would be close to zero, potentially underestimating the user's sewing expertise.

Overall, our prototype implements a digital version of the workflow we discovered during our interviews: Students practice sewing by stitching on or along lines printed on a sheet of paper or other material. Our software demonstrates that this sewing expertise metric can be used to implement automated expertise assessment. The accuracy of the app is limited by the OpenCV image recognition we use and the optimized overlapping alignment of the *ICP* algorithm [4, 28]. To reduce computation time and provide quick feedback, our prototype currently tests overlapping alignment only for the center point of the target line and assumes that the two lines are not rotated against each other. We omit rotation because our task requires users to sew at a fixed distance, which is represented by a rotation-free constant translation of the target line. Testing further or even all pixel points of each line and including rotation could increase the accuracy of the algorithm at the cost of rapid feedback.

The code of our demonstrator is available at https://hci.rwth-aachen.de/sewing-assessment.

## 7 CONCLUSION & FUTURE WORK

Since it is difficult for makers to get feedback on their skills, we see potential in automated systems that can support the assessment process and lower its inherent bias. However, such systems require expertise-related metrics for the assessment, which we identified for the sewing domain (Tab. 2 & 3).

These metrics were derived from an interview study with ten professional sewing-related instructors who shared how they evaluate their students. To showcase the applicability of our metrics, we presented a demonstrator prototype that assesses the underlying properties of one such metric, the *seam allowance*. It evaluates the ability to sew along a constant distance to a target line, which is a required skill for sewing a good *seam allowance*.

Systems that utilize our reported metrics could, for example, analyze materials and techniques used in previous sewing projects of a user. Such systems could suggest suitable projects that use materials with properties similar to those used in completed projects, potentially leading to a gradual improvement in sewing skills. This vision is based on the insights from our interviewees that different fabrics and techniques require different levels of expertise.

[3]https://developer.apple.com/documentation/coreimage/cidetectortyperectangle (accessed Mar 22, 2024)
[4]https://opencv.org (accessed Mar 22, 2024)
[5]https://www.bernina.com/en-US/Machines-US/Series-Overview/BERNINA-8-Series (accessed Mar 22, 2024)

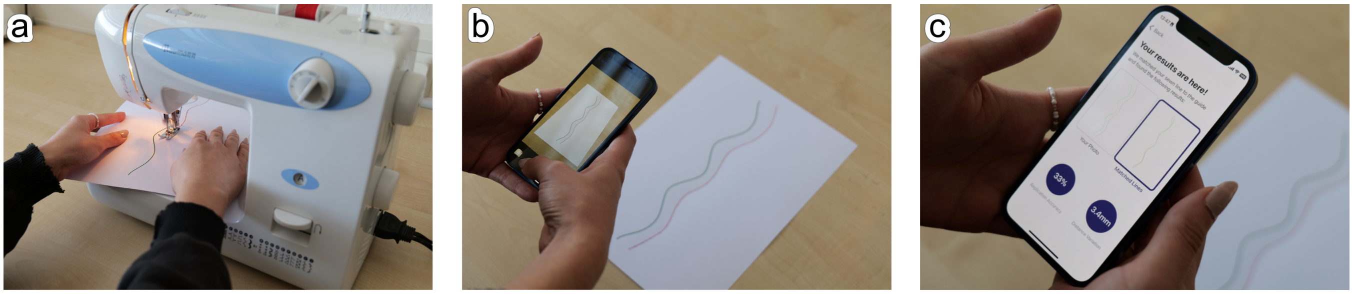


**Figure 2: Using our prototype to assess *seam allowance* quality. The user sews onto a printed worksheet, aiming to stay at a constant distance from the given line (a), photographs the result (b), and the app displays the detection and the evaluation results (c).**

We aim to continue this work by building upon our demonstrator. It utilizes established algorithms to digitize a workflow that according to our interviewees is widely used. We also intend to investigate whether this digital approach creates new possibilities to support the sewing learning process. Further, we want to create additional sewing-expertise assessment artifacts based on the other metrics we identified. Ultimately, such systems may not only be able to assist in self-assessment but also support instructors like our interviewees in their teaching.

On a broader scale, we expect automated crafting expertise-assessment systems to support users in practicing their skills, mitigating the effects of self-assessment bias. Further, identifying metrics in other crafting domains can advance research on this topic and might discover analogies between metrics of different domains to develop a more general crafting expertise assessment model.

## REFERENCES

[1] Michelle Annett, Tovi Grossman, Daniel Wigdor, and George Fitzmaurice. 2019. Exploring and Understanding the Role of Workshop Environments in Personal Fabrication Processes. *ACM Transactions on Computer-Human Interaction* 26, 2, Article 10 (mar 2019), 43 pages. https://doi.org/10.1145/3301420

[2] Christiane Attig, Ester Then, and Josef F. Krems. 2019. Show Me How You Click, and I'll Tell You What You Can: Predicting User Competence and Performance by Mouse Interaction Parameters. In *Intelligent Human Systems Integration 2019*, Waldemar Karwowski and Tareq Ahram (Eds.). Springer International Publishing, Cham, 801–806.

[3] Marc Bächlin, Kilian Förster, and Gerhard Tröster. 2009. SwimMaster: A Wearable Assistant for Swimmer. In *Proceedings of the 11th International Conference on Ubiquitous Computing (UbiComp '09)*. Association for Computing Machinery, New York, NY, USA, 215–224. https://doi.org/10.1145/1620545.1620578

[4] P.J. Besl and Neil D. McKay. 1992. A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14, 2 (1992), 239–256. https://doi.org/10.1109/34.121791

[5] David Buckley, Ke Chen, and Joshua Knowles. 2017. Rapid Skill Capture in a First-Person Shooter. *IEEE Transactions on Computational Intelligence and AI in Games* 9 (2017), 63–75. https://doi.org/10.1109/TCIAIG.2015.2494849

[6] David N. Chin. 1989. KNOME: Modeling What the User Knows in UC. In *User Models in Dialog Systems*, Alfred Kobsa and Wolfgang Wahlster (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 74–107.

[7] Victoria Clarke and Virginia Braun. 2017. Thematic Analysis. *The Journal of Positive Psychology* 12, 3 (2017), 297–298. https://doi.org/10.1080/17439760.2016.1262613 arXiv:https://doi.org/10.1080/17439760.2016.1262613

[8] Michel C. Desmarais and Jiming Liu. 1993. Exploring the Applications of User-Expertise Assessment for Intelligent Interfaces. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems (CHI '93)*. Association for Computing Machinery, New York, NY, USA, 308–313. https://doi.org/10.1145/169059.169243

[9] Arin Ghazarian and S Majid Noorhosseini. 2010. Automatic detection of users' skill levels using high-frequency user interface events. *User Modeling and User-Adapted Interaction* 20 (2010), 109–146.

[10] Jun Gong, Fraser Anderson, George Fitzmaurice, and Tovi Grossman. 2019. Instrumenting and Analyzing Fabrication Activities, Users, and Expertise. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300554

[11] Jodi S Goodman. 1998. The Interactive Effects of Task and External Feedback on Practice Performance and Learning. *Organizational Behavior and Human Decision Processes* 76, 3 (1998), 223–252.

[12] Robert D Grober. 2010. An Accelerometer Based Instrumentation of the Golf Club: Measurement and Signal Analysis.

[13] Tovi Grossman and George Fitzmaurice. 2015. An Investigation of Metrics for the In Situ Detection of Software Expertise. *Human–Computer Interaction* 30 (2015), 64–102. https://doi.org/10.1080/07370024.2014.881668

[14] Pamela J Hinds. 1999. The Curse of Expertise: The Effects of Expertise and Debiasing Methods on Predictions of Novice Performance. *Journal of Experimental Psychology: Applied* 5 (1999), 205.

[15] Pamela J Hinds, Michael Patterson, and Jeffrey Pfeffer. 2001. Bothered by Abstraction: The Effect of Expertise on Knowledge Transfer and Subsequent Novice Performance. *Journal of Applied Psychology* 86 (2001), 1232.

[16] Amy Hurst, Scott E. Hudson, and Jennifer Mankoff. 2007. Dynamic Detection of Novice vs. Skilled Use Without a Task Model. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. Association for Computing Machinery, New York, NY, USA, 271–280. https://doi.org/10.1145/1240624.1240669

[17] Aftab Khan, Sebastian Mellor, Eugen Berlin, Robin Thompson, Roisin McNaney, Patrick Olivier, and Thomas Plötz. 2015. Beyond Activity Recognition: Skill Assessment from Accelerometer Data. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 1155–1166. https://doi.org/10.1145/2750858.2807534

[18] Stacey Kuznetsov and Eric Paulos. 2010. Rise of the Expert Amateur: DIY Projects, Communities, and Cultures. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries (NordiCHI '10)*. Association for Computing Machinery, New York, NY, USA, 295–304. https://doi.org/10.1145/1868914.1868950

[19] Cassim Ladha, Nils Y. Hammerla, Patrick Olivier, and Thomas Plötz. 2013. ClimbAX: Skill Assessment for Climbing Enthusiasts. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing (UbiComp '13)*. Association for Computing Machinery, New York, NY, USA, 235–244. https://doi.org/10.1145/2493432.2493492

[20] Mackenzie Leake, Kathryn Jin, Abe Davis, and Stefanie Mueller. 2023. InStitches: Augmenting Sewing Patterns with Personalized Material-Efficient Practice. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 267, 14 pages. https://doi.org/10.1145/3544548.3581499

[21] Edwin A Locke and Gary P Latham. 2002. Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey. *American Psychologist* 57 (2002), 705.

[22] Richard E Mayer and Roxana Moreno. 2003. Nine Ways to Reduce Cognitive Load in Multimedia Leaming. *Educational Psychologist* 38 (2003), 43–52.

[23] Richard E Nisbett and Timothy D Wilson. 1977. Telling More Than We Can Know: Verbal Reports on Mental Processes. *Psychological Review* 84 (1977), 231.

[24] Donald A Norman. 2013. *The Design of Everyday Things: Revised and expanded edition.* Basic Books, New York City, USA.

[25] Fred Paas, Alexander Renkl, and John Sweller. 2003. Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist* 38 (2003), 1–4.

[26] C Richards, J Rosen, B Hannaford, C Pellegrini, and M Sinanan. 2000. Skills evaluation in minimally invasive surgery using force/torque signatures. *Surgical Endoscopy* 14 (2000), 791–798.

[27] J. Rosen, B. Hannaford, C.G. Richards, and M.N. Sinanan. 2001. Markov Modeling of Minimally Invasive Surgery Based on Tool/Tissue Interaction and Force/Torque Signatures for Evaluating Surgical Skills. *IEEE Transactions on Biomedical Engineering* 48 (2001), 579–591. https://doi.org/10.1109/10.918597

[28] S. Rusinkiewicz and M. Levoy. 2001. Efficient Variants of the ICP algorithm. In *Proceedings Third International Conference on 3-D Digital Imaging and Modeling*. IEEE, Piscataway, USA, 145–152. https://doi.org/10.1109/IM.2001.924423

[29] Shahed Anzarus Sabab, Adnan Khan, Parmit K. Chilana, Joanna McGrenere, and Andrea Bunt. 2020. An Automated Approach to Assessing an Application Tutorial's Difficulty. In *2020 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. IEEE, Piscataway, USA, 1–10. https://doi.org/10.1109/VL/HCC50065.2020.9127271

[30] Johnny Saldaña. 2013. *The Coding Manual for Qualitative Researchers*. sage, London.

[31] Nick Taylor, Ursula Hurley, and Philip Connolly. 2016. Making Community: The Wider Role of Makerspaces in Public Life. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1415–1425. https://doi.org/10.1145/2858036.2858073

[32] Cristen Torrey, Elizabeth F. Churchill, and David W. McDonald. 2009. Learning How: The Search for Craft Knowledge on the Internet. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. Association for Computing Machinery, New York, NY, USA, 1371–1380. https://doi.org/10.1145/1518701.1518908

[33] Marco Wolf and Shaun McQuitty. 2011. Understanding the do-it-yourself consumer: DIY motivations and outcomes. *AMS Review* 1 (2011), 154–170.

[34] Gabriele Wulf, Charles Shea, and Rebecca Lewthwaite. 2010. Motor skill learning and performance: A review of influential factors. *Medical education* 44, 1 (2010), 75–84.

[35] Gabriele Wulf, Charles H Shea, and Sabine Matschiner. 1998. Frequent Feedback Enhances Complex Motor Skill Learning. *Journal of Motor Behavior* 30, 2 (1998), 180–192.