Tacit Descriptions: Uncovering Ambiguity in Crowdsourced Descriptions of Motions and Materials

Shreyosi Endow
shreyosi.endow@mavs.uta.edu
The University of Texas at Arlington
Arlington, Texas, USA

Cesar Torres
cearto@uta.edu
The University of Texas at Arlington
Arlington, Texas, USA

ABSTRACT

Although a picture, GIF, or tutorial video is "worth a thousand words", every viewer has a unique perspective of which words they 'see.' Packed with rich meaning, these words, or descriptions, are often used to describe the ambiguity of working with materials, artifacts, and processes in creative practices. Our work reframes the traditional crowdsourced description tasks to leverage crowdworker diversity in generating a corpus of tacit descriptions. These descriptions are synthesized into a typology and describe how users communicate ambiguity and tacit knowledge embedded in unfamiliar material properties and familiar handed motions. We propose three design concepts that demonstrate how tacit descriptions could enhance the interpretation of tutorial artifacts - Tacit Layer, or breadth-first interpretation, Tacit Space, or depth-first interpretation, and Tacit Localizer, or context-based interpretation. These concepts are then translated into physical card-sort activities and used in a workshop to gauge how participants consider tacit information in their respective practices.

CCS CONCEPTS

• **Human-centered computing** **Interaction design process and methods**; *Empirical studies in interaction design*.

KEYWORDS

tutorials; tacit knowledge; creative crowdsourcing; typology; design research

ACM Reference Format:
Shreyosi Endow and Cesar Torres. 2023. Tacit Descriptions: Uncovering Ambiguity in Crowdsourced Descriptions of Motions and Materials. In *Designing Interactive Systems Conference (DIS '23), July 10–14, 2023, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3563657.3596048

1 INTRODUCTION

> *Language struggles in depicting physical action, and nowhere is this struggle more evident than in language that tells us what to do* - Richard Sennett

The tutorial is a central vehicle for how creative practices communicate knowledge. Yet when attempting to communicate *tacit knowledge*, or "things that we know but cannot tell" [30], tutorials are fundamentally limited by the differences in materials, tools, skill set, and mental models of the tutorial author and tutorial user. Whether executing a recipe, bio-protocol, scientific experiment, or fabrication process, HCI researchers have sought to "fill-in-the-gap" by improving the information bandwidth of tutorials through interactive media [9, 28, 40]. And although pictures, GIFs, videos, or creative live streams continue to add to the axiomatic "thousand words" transfer, there remains a need to acknowledge how each tutorial user has a unique lens that filters which "words" they choose to see.

Much like an eye tracker is used to document visual attention, we present a crowd-based method for tracking which visual information is being cognitively interpreted. Akin to the traditional crowd labeling tasks, our method leverages the diversity of crowdworkers [15, 16, 27] to annotate short tutorial GIFs with semantic descriptions.

In this work, we describe one investigation aimed at understanding how interpretation is affected when viewing familiar and unfamiliar experiences. Deployed as an Amazon Mechanical Turk HIT[1], we collected over 600 descriptions from over 70 participants on two open-ended description tasks which asked participants to describe whisking motions (*familiar experience*) and clay consistencies (*unfamiliar experience*) as depicted in GIFs. Using an inductive coding process, we synthesize these descriptions into a typology called the Tacit Description Typology (TDT) that organizes the different ways users semantically communicate explicit and tacit information. We propose a set of design concepts grounded in the typology that demonstrate how tacit descriptions can be used to enhance the reading of tutorial artifacts. They include (1) Tacit Layer, which injects relevant information into a user's view to provide a comprehensive understanding of the task at hand, (2) Tacit Space, which creates and demarcates the boundaries between positive and negative actions for a specific task, and (3) Tacit Localizer, which aids a user in "finding" themselves within a task or practice by matching them to other users who see, think, or talk like them.

To ground our thinking of tacit descriptions, we translate our design concepts into a set of physical card-sort activities, which are then used in a workshop with six participants to engage, provoke, and gauge how participants consider tacit information and envision tutorial interactions in their respective practices.

Our work takes a step towards the design and development of tutorial systems that identify and bridge the gap in tacit knowledge communication. We contribute guidelines for sourcing tacit descriptions from the crowd, a description typology to gauge label

[1]A Human Intelligence Task, or HIT, is a job, or microtask, posted by a requester on Amazon Mechanical Turk for a crowdworker to complete for compensation [15, 27].

DIS '23, July 10–14, 2023, Pittsburgh, PA, USA

ACM ISBN 978-1-4503-9893-0/23/07.
https://doi.org/10.1145/3563657.3596048

saturation, typology-driven card sort activities for design ideation in a workshop setting, and three design concepts for advancing tutorial systems.

The paper is organized to first describe how the design community sources and uses tacit information (Section 2). We provide an overview of our methods (Section 3), then describe our crowdsourcing technique and coding process for synthesizing tacit descriptions into the Tacit Description Typology (TDT) (Sections 4 and 5). We reflect on the replicability and scalability of our crowdsourcing method and provide guidelines for other researchers who may want to follow a similar method (Section 6). We then describe the three design concepts that were derived from TDT (Section 7) and walkthrough how the concepts were translated into card-sorting activities for the workshop (Section 8). We then present our findings from the workshop on how people consider tacit information in tutorial interactions (Section 8) and finally, discuss design implications for leveraging tacit descriptions in new tutorial mediums and psychometric scales (Section 9).

2 RELATED WORK

Tacit information and communication is an interdisciplinary area of research that draws from psychology, human-computer interaction, and design communities. We position our work with regard to ongoing efforts to communicate and transfer tacit information. We describe the role of tacit information within the tutorial and instructional materials, methods for extracting tacit information, and, more specifically, efforts in extracting descriptions from GIF media.

2.1 Tacit Communication through Tutorials

Tacit information refers to knowledge that is difficult to relay in words [30]. In many situations, tacit information becomes subconscious [23], making it difficult for a tutorial author to remember what it was like not to know [32]. Even if an author may not be aware of what they are doing, Edwards et al. found that skilled practitioners would 'deny vigorously' carrying out a task in a certain way in spite of video documentation [8]. When creating tutorial content, this inability or reluctance to access tacit knowledge by tutorial authors often results in an omission of valuable and experiential details that could have otherwise benefited inexperienced learners [8].

Sennett [32] advocated for moving away from over-informative, dry, and technical descriptions towards three types of expressive instructions capable of relaying tacit knowledge. First, *sympathetic illustration*, where the expert or author "emotionally returns" to their novice days, relays instructions from the emotional viewpoint of the learner. Second, *scene narrative* reframes the instructional task as part of a larger narrative, or process, to decenter the learner and enable them to imagine the whole process while providing advice indirectly. Third, *instruction through metaphors* sources the language from metaphors to "give each physical action heavy symbolic weight" and allow the learner to form a more meaningful understanding of their actions. These strategies resonate with existing research that highlights the growing importance of creating metaphors to represent the relationship between theory and practice [2]. Cognitive aids such as mnemonics, similes, metaphors, and idioms are often used in communicating tacit knowledge because of their ability to make abstract and complex concepts more concrete, memorable, and relatable. Others have proposed the need to move beyond semantics and engage with other sensemaking capabilities. In an analysis of tutorial mediums, Endow et al. [9] described the need for tutorial systems that redirect feedback to the perceptual channels of the human body and foreground a user's ability to perceive and make sense of their own activity. Smith et al. [33] leveraged data visceralization to sonify the movements of users engaged in crocheting activities, which were found to help practitioners reflect upon each other's actions. Our work offers an approach where tacit information is purposefully elicited from the crowd as opposed to the author to obtain the obvious information needed to fill in the gap in instructions; as a crowd task, we leverage the natural bias in users to attend to different perceptual stimuli to extract descriptions that span the sensemaking modalities of the body.

2.2 Tutorial Design and Authoring

Intelligent tutorial systems aim to find the happy medium between too much and too little information. A core strategy has been improving existing tutorial content. Digestibility approaches extract structured information using hierarchical [35], domain-specific [22], and feature-based [6] segmentation. Context-matching approaches instead aim to match the user's situation and need with available information through techniques like orientation mapping [38] or retargeting information to contextual form factors [28, 40]. However, the existing corpus of tutorials is skewed toward English speakers and continues to marginalize tutorial audiences, especially when leveraging NLP techniques [11]. In contrast to video and written instructions, Jones et al. [17] leveraged craft pattern practices to relay information as tangible microtasks that foreground the material experience. Our work identifies the range of descriptions that are used to communicate tacit activities and offers a typology for improving the interpretability of instructions.

Improving the tutorial *authoring* experience has been limited to domain-specific accelerators that streamline the tutorial authoring process by making relevant information such as spatial cues [4] or tool interactions [22] easier to capture and annotate. Community-enhanced tutorials [21] can improve authored content through community contribution, but this approach suffers from the need to curate community contributions for quality and similarity. We follow a similar approach where tutorial content (tacit descriptions) is sourced from the crowd; unlike other crowd labeling approaches that aim to refine labels, our approach leverages crowd diversity as a means to unpack the tacit knowledge that is unknowingly encoded in video instructions.

2.3 Extracting Tacit Information

Since tacitness is a property of the knower, extracting and synthesizing tacit knowledge often relies on understanding the differences in knowledge between the expert and the novice [37]. Wood et al. [36] proposed an expert learner approach where an individual skilled in a peripheral area documents their skill development under the tutelage of a skilled practitioner. Torres et al. [34] proposed analyzing sensor logs that documented the minute-by-minute changes in

Figure 1: A. Static frames of clay GIFs. 1) Rocks in a bucket. 2) Reclaiming clay bucket. 3) Initial wedge consistency. 4) Final wedge consistency. 5) Leather hard clay. B. Static frames of whisking GIFs. 1) Jostling milk. 2) Whipping cream. 3) Incorporating cream. 4) Beating eggs. 5) Mixing vinaigrette.

skill development in repeated task experiments. Within material-centered design practices, a regular finding is that tacit information is situated between the material and user; techniques like sensory augmentation have shown promise in relaying tacit information within ceramics practices [14]. Ross et al. encountered that creative thought, and by extension making use of tacit knowledge, requires action [31]. Our approach leverages the different expertise and skill levels of crowdworkers; although crowdworkers in our study do not interact directly with a material, we find that they are able to project themselves into our GIF tasks and describe information that is not present (e.g., sound). We find that verbalizing their interpretation of a GIF has value as a form of enaction.

2.4 Generating Descriptions from GIFs

Generating descriptions for accessibility is a growing research area within machine learning and crowdsourcing. Zhang et al. [39] found that semistructured tasks that outline qualities to include in an annotation were the most successful for generating high-quality information for Blind and Low Vision users (BLV). Description tasks have, however, been focused on relaying the rote visual content – the TGIF dataset showed that ML models were capable of annotating GIFs with near accuracy to crowd-supplied annotations [24]; however, the gap between human and machine annotation appears to remove subjective and tacit details: "a whale is swimming the ocean (machine)" + "slowly (human)"; "a ballerina is performing a dance routine (machine)": "ballerina performs a beautiful dance with a smile (human)". Silva et al. [25] found that NLP-generated labels did not match the richness with which users described their food intake within a food journaling application. These annotations used contextual cues (like location or social circumstances, e.g. "dinner at friend's house, bbq chicken with mac and cheese") and specified amounts of food consumed using subjective references that made personal sense (e.g. ".25 pkg Trader Joe's Asian noodle salad[...]"), yet were uninterpretable by commercial NLP systems. Our approach explores open-ended annotations to capture both interpretation (reading a GIF) and communication (articulating the interpretation semantically).

3 METHODS

To understand how people describe physical situations where tacit knowledge is present, we first conducted a study with two open-ended description tasks which anticipate sentence-like or multi-word responses. Both tasks asked respondents to articulate what they see in a visual prompt in a GIF format through a written description. Our goal was to gather these tacit descriptions to understand patterns in how people chose to perceive or interpret the visual prompt and how they relayed this information back semantically. We synthesized the tacit descriptions using an inductive coding process into a typology called the Tacit Description Typology (TDT).

As a demonstration of how TDT can be leveraged in designing systems where tacit knowledge is prevalent, we created three design concepts grounded in exemplar scenarios - Tacit Layer, Tacit Space, and Tacit Localizer. To further ground our thinking of tacit descriptions, we translated these design concepts into a set of physical card-sort activities that were presented to six participants in a workshop study. Through the workshop study, we gauged how participants envisioned our design concepts in tutorial interactions in their respective practices. We conducted a thematic analysis [3, 7] on video recordings of participants' interviews and card activities from the workshop and consolidated our findings into descriptive themes.

4 EXTRACTING AND UNDERSTANDING TACIT DESCRIPTIONS

In this section, we first describe the two open-ended description tasks which were collectively deployed to participants in one survey and our rationale behind the task choice. We then walk through our data collection, cleaning, and coding processes.

4.1 Open-Ended Description Tasks

The first task examines the language and understanding used by individuals when describing an unfamiliar material without actively interacting with it. The second task explores the aspects of a familiar motion that are most salient to people when describing the action from an observer's perspective.

Task 1: Describing Clay Consistencies. The first task involved articulating the consistencies of five different states of clay. Working

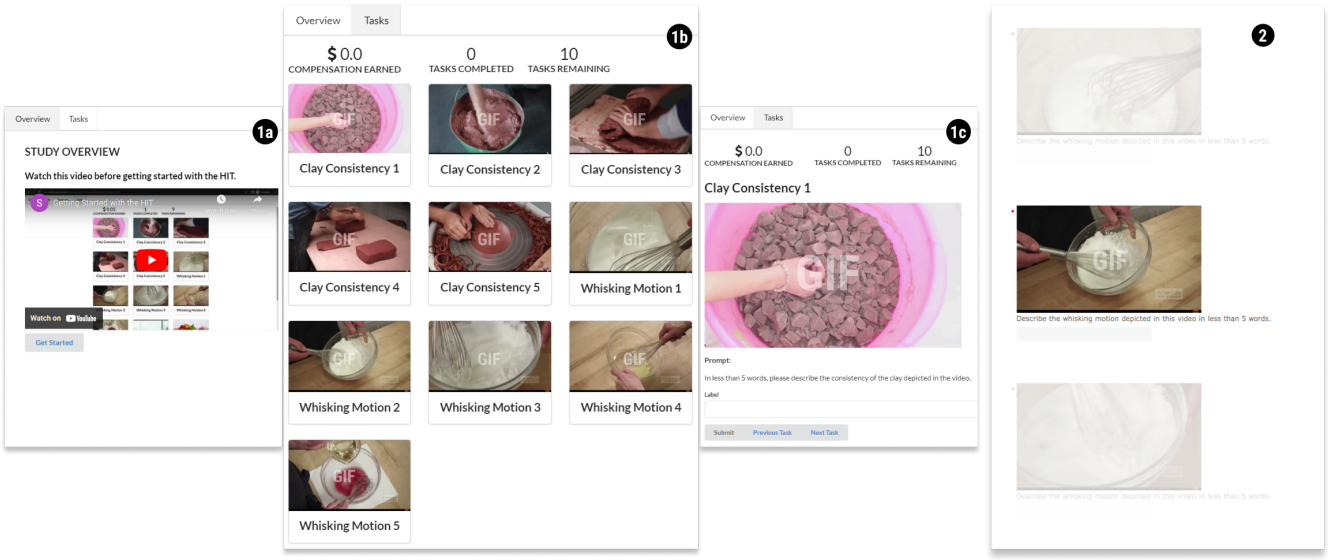

Figure 2: 1. Web-based application for extracting responses. a) The landing page provides an overview of the study and examples of the types of descriptions we are looking for through a tutorial. b) The tasks page showcases all the GIFs in a card view as well as helps the respondent keep track of their progress. c) The individual task page shows a larger version of the GIF and prompts the respondent for a response. 2. QuestionPro survey for extracting responses. A segment of the QuestionPro survey showing the GIFs and prompts embedded in the survey itself.

with clay is a practice that is rich in unspoken, subconscious, and tacit actions [9]; the consistency of clay varies greatly depending on its environmental conditions and functional requirements and is visually, haptically, and aurally distinguishable. Despite encountering ceramic artifacts day-to-day, people outside the ceramics community are often unfamiliar with the various states of clay involved in creating ceramic artifacts. Our choice of using clay consistency was further motivated by the tacit descriptions commonly used by clay practitioners to define clay states. Typically, clay practitioners draw comparisons to other materials (bone dry, leather hard) or describe the consistency according to the process it is prepared for (wedging consistency, slip consistency). We were interested in exploring if this vocabulary extended to a larger, non-expert crowd and, if not, what material qualities or properties were being tapped into to describe the consistency.

Task 2: Describing Whisking Motion. The second task involved articulating a whisking motion. As a staple of cooking and baking experiences, hand whisking, whether it involves mixing, beating, or stirring ingredients together, does not require niche expertise and is familiar to a larger audience. The familiarity of the task also invites ambiguity as people may not necessarily be explicitly aware of their specific actions or preferences for the way they use a whisk.

Task Breakdown. Short videos(<30 s) of five different clay consistencies were extracted from two popular YouTube videos on manipulating clay [12, 26]. Similarly, we extracted five (< 30 s) videos of different whisking motions from a popular whisking tutorial video on YouTube [19]. These clips were all rendered into GIFs. Figure 1 shows a static frame of the GIFs; full GIFs are provided as supplemental material. The following prompt was provided for each clay GIF: *In less than 5 words, please describe the consistency of the clay depicted in the video.* The following prompt was provided for each whisking GIF: *In less than 5 words, please describe the whisking motion depicted in the video.* We kept the prompt open-ended to ensure that respondents' descriptions reflect their natural and unbiased vocabulary and understanding. Additionally, we set the word count to be less than 5 words to appeal to crowdworker preferences for shorter tasks [29].

4.2 Data Collection and Analysis

To reach a varied sample of people [15, 16, 27], we conducted the study on Amazon Mechanical Turk, a crowdsourcing platform that is popular for survey-based studies among HCI researchers. Both tasks were consolidated in one survey, which consisted of 10 subtasks (5 clay prompts and 5 whisking prompts). The survey was deployed in two formats - 1) a traditional QuestionPro survey and 2) a custom web-based application. A walkthrough of the web app is shown in Figure 2 along with a snippet of the QuestionPro survey. In addition to collecting descriptions, the survey also asked respondents for demographic information.

Data Collection. We collected descriptions from 63 respondents using the QuestionPro format and 10 respondents using the web application. We pooled the responses from the two formats (total 73 respondents) since we observed no visible differences in data quality.

Data Cleaning. To improve the internal validity of the data collected, we removed responses that met three different criteria.

Duplicate Responses We removed all responses from 5 respondents who provided the same description for two or more GIFs which reduced the overall respondent pool to 68 respondents.

Partial Survey Completion Although the survey consisted of both clay and whisking GIFs, some respondents provided descriptions for only one category of GIFs (clay or whisking). In particular, 4 respondents did not provide descriptions for the clay tasks, reducing the clay respondent pool to 64 respondents, and 1 respondent did not provide responses for the whisking tasks, reducing the whisking respondent pool to 67 respondents.

Poor descriptions We further removed 26 responses from the clay tasks and 20 responses from the whisking tasks for being what we termed poor descriptions. These descriptions either did not relate to the GIFs or were nonsensical. Often, these responses seemed to be generated by a text generator or AI and did not resemble the task at all.

The remaining descriptions were cleaned to correct minor spelling mistakes.

The final dataset for clay descriptions comprised 294 descriptions or data points (73 total respondents - 5 respondents with duplicate responses - 4 respondents with partial survey completion = 64 respondents x 5 GIFs = 320 descriptions - 26 poor descriptions = 294 descriptions) from 64 respondents (Average age 35 ± 0.88 years, 37 Male, 25 Female, 2 Other).

The final dataset for whisking descriptions comprised 315 data points (73 total respondents - 5 respondents with duplicate responses - 1 respondent with partial survey completion = 67 respondents x 5 GIFs = 335 descriptions - 20 poor descriptions = 315 descriptions) from 67 respondents (Average age 35 ± 0.86 years, 38 Male, 27 Female, 1 Non-binary, 1 preferred not to say).

Coding. All cleaned descriptions were recorded on a shared Google Sheet between two paper authors who independently conducted an inductive coding process to avoid bias in code generation. Each word in a description was individually coded. We additionally performed keyword matching using spreadsheet macros to ensure that each word has the same code across all descriptions. To streamline the process of reaching a code agreement, we implemented a Python script that read all the descriptions and their respective codes from the two separate authors and generated a new sheet consisting of all descriptions with mismatched codes. The two authors reevaluated their codes for these descriptions and updated the codes until at least 90% agreement was reached. Both authors then collaborated to cluster the codes into subtypes and identified common themes among subtypes. Subtypes were then further clustered based on the common themes and each subtype cluster was assigned a representative type name to create the final typology.

5 THE TACIT DESCRIPTION TYPOLOGY (TDT)

From our coding process, we derived a total of 12 description types from both tasks; the distributions of these description types across all descriptions are shown in Figure 4. Our Tacit Description Typology (TDT) consolidates these 12 description types into four larger types - INSTRUCTION, EXPERIENCE, METAPHOR, and CARE. Several descriptions spanned multiple types of the typology, demonstrating people's tendencies to combine different sensory, experiential, and tacit information to create more complete, coherent, and accurate perceptions of materials and motions. Figure 3 shows the types and subtypes along with their definitions and example descriptions. We elaborate on the four types - INSTRUCTION, EXPERIENCE, METAPHOR, and CARE below.

5.1 INSTRUCTION

The INSTRUCTION type encompasses the ACTION, DIRECTION, PACE, MOTION, PROCESS, and STOP CONDITION subtypes.

The instruction type is focused on capturing descriptions of motions or materials that closely resemble those found in instructional materials. In particular, these descriptions focus on explicit knowledge, contain both technical and functional qualities of the motion and the material, and are informed by interactions.

For instance, respondents described the consistency of clay with respect to its functional capabilities (e.g. malleability, stretchability), shedding light on the close tie between our perception and interactions [ACTION]. We see this among clay practitioners who describe the material with reference to what can be done with it. For example, a "leather hard" clay [consistency] describes a state where clay can be trimmed without risking the clay from collapsing on itself [ACTION]. When it came to describing the whisking motion, respondents primarily fixated on the relative movement [DIRECTION] and speed [PACE] of the whisk in the depicted environment. Direction descriptions leveraged both linear directional language like "back and forth", as well as shape-based directions such as swirls, ovals, or circles to describe the motion. However, rarely was a direction communicated in a scientific or objective manner that specified the exact displacement in a formal unit such as inches or centimeters. Similarly, the pace of the motion was described using subjective and colloquial terms such as fast, rapid, and brisk, which are inherently ambiguous and can vary greatly based on individual perceptions and experiences. Only one respondent described the whisking speed objectively as revolutions per second.

This is similar to how tutorial authors often describe qualities or quantities in experiential and informal ways. For example, in a cooking tutorial, the author might describe the quantity of salt needed as "a dash" or "a pinch" rather than explicitly specifying "a teaspoon" of salt.

Many respondents found it more efficient to describe whisking using domain-specific technical terms such as "beating", "whipping", "incorporating", and "folding", which capture the direction, intensity, and hand positioning, rather than breaking down the whisking motion into its various components [MOTION]. These terms can potentially convey a sense of expertise and precision, for example, "julienning" vegetables convey a finer and more controlled cut in comparison to "chopping" vegetables.

Some respondents considered the whisking motion as part of a larger process [PROCESS], highlighting how often the motion should be repeated and when to stop [STOP CONDITION], and resembling the flow of information in instructional materials. These descriptions often focused on the state of the material once the whisking was done, even though this information was not provided in the GIFs. Many of these descriptions can be attributed to prior experiences with tasks that are commonly associated with whisking,

TYPE	SUBTYPE	SUBTYPE DEFINITION	EXAMPLES
EXPERIENCE	Haptics	Characterizes the material's tactile properties	**hard** and rocky [hardness] **soft** and clumped [softness]
	Visuo-haptics	Integrates both visual and haptic modalities	**smooth** [texture] and **dense** [form] **thick** [shape] and **heavy** [weight]
	Visual	Uses only the visual sensing modality to describe the appearance	**glaze** and **glossy** [look] **choppy** [look]
	Sound	Uses a sound word to describe the material or motion	rough and **crunchy** **whispy**
INSTRUCTION	Direction	Characterizes relative displacement in an arbitrary space	**figure eights** [softness] careful **swirls** [shape]
	Motion	Uses a technical word to capture multiple aspects of movement	careful **blending** [common] **aerating** whip [advanced]
	Pace	Describes the speed at which the whisk is being used	**briskly** back and forth **slow** and **moderate**
	Process	Describes the motion as part of a process	fast **repetitive** sideways movement [rhythm] **constant** agitation [duration]
	Stop Condition	Uses the end state of the material or process as a way of describing motion	stir it **until creamy** [material] forceful and **complete** [process]
	Action	Functional descriptions of clay consistencies informed by interactions	hard and **peelable** [interaction] **malleable** [property] and slightly moist
METAPHOR	Metaphor	Draws a comparison with another material or motion	**tofu** firm consistency [food] **bitch slapping motion**
CARE	Care	Refers to the level of care, skill, attitude and intention behind the motion	**steady** [skill] and **purposeful** [attitude] **deliberate** but messy swirls [intention]

ONLY WHISK ONLY CLAY BOTH WHISK AND CLAY

Figure 3: Tacit Description Typology (TDT). Overview of the typology types and subtypes with example descriptions pertaining to each subtype. The words in the descriptions that correspond to each subtype are bolded, with more granular codes provided in parenthesis. A clay carving tool icon and a whisk icon indicate subtypes that were present in the clay and whisking tasks respectively. An icon combining the carving tool and whisk indicates subtypes that were present in descriptions from both tasks.

such as beating or whisking egg whites or cream until soft peaks form - a common use of a handheld whisk.

Overall, the instruction type encompasses descriptions that leverage domain-specific technical terms to describe functional qualities of motions and materials similar to those found in instructional materials.

5.2 EXPERIENCE

The EXPERIENCE type encompasses the HAPTICS, VISUO-HAPTICS, VISUAL, and SOUND subtypes.

The experience type comprises descriptions that focus on the sensory experiences of perceiving or interacting with a material or a motion. In particular, many respondents tapped into their senses of touch to describe the material [HAPTICS], characterizing the softness, hardness, or sharpness of the material, which are perceived by mechanoreceptors on the skin. The ability to associate a tactile sensation with a purely visual input is informed by prior experiences with other materials; even if respondents never felt the specific consistencies of clay depicted in the GIFs, they have interacted with materials that appear similar.

Descriptions also isolated visual characteristics of the material without describing any other sensory information [VISUAL]. These descriptions spoke of the appearance of the material in terms of its consistency as well as the distribution of the material in the larger workspace.

We saw a large number of descriptions that combined both senses of sight and touch to shed light on the texture, shape, and size of the clay [VISUO-HAPTICS]. Our ability to perceive texture, for

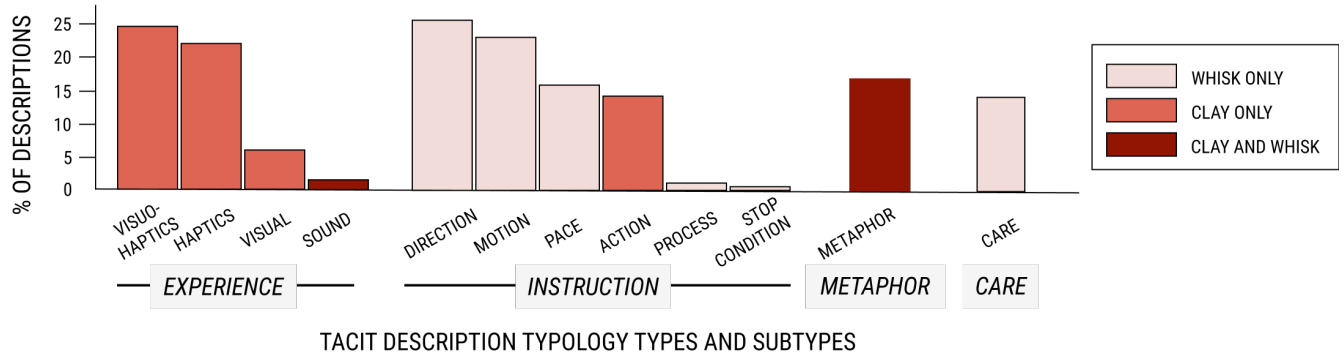

Figure 4: Percentage of TDT subtypes across all collected descriptions.

example, involves the simultaneous processing of the visual characteristics of the material, such as the surface features, and associating the visual characteristics with specific tactile sensations based on our past experiences.

Although rare, some respondents described the motion and the material consistency with phonetics or onomatopoeia [SOUND]. These descriptions were particularly interesting as the GIFs had no sound, indicating that these descriptions purely stemmed from memory and prior experiences.

5.3 METAPHOR

The METAPHOR type consolidates descriptions that liken the material or the motion to another through metaphoric language and specifically stem from experiences. These descriptions belong to the larger METAPHOR type without specific subtypes as there were not enough distinguishable themes among the descriptions to justify clustering at a subtype level. Descriptions in this category are some of the most information-dense as the metaphors capture nuances from the other typology subtypes as well.

Contrary to the MOTION code in the instruction type, descriptions in this type not only use technical motion terminology but also demonstrate creativity and richness in the comparisons. Additionally, the metaphors extend beyond the motion to the material being whisked despite the prompt asking explicitly to describe the motion, indicating that for some people, the motion cannot be distinguished from the material being manipulated. A few descriptions also leverage idioms to describe the whisking motion in a figurative manner, highlighting the importance of understanding cultural and linguistic context when interpreting a motion - someone who is not familiar with colloquial English terminology may not correctly interpret what whisking "all over the place" means.

When describing the clay consistency, several respondents drew parallels to food (eg. pudding, whipping cream, tofu, dough, meat), as the tactile and visual characteristics of the material resembled those of certain food items despite its in-edibility. The texture, shape, and consistency of the clay may trigger memories and associations with food items, enabling respondents to relate to the clay in a more relatable and familiar way.

5.4 CARE

The CARE type comprises of descriptions that assess the level of care, attention, intent, attitude, and skill behind the interactions with the motion or material. The type name is inspired by Sennett [32] who defines care as the distinguishing quality between a craftsman and an amateur; the craftsman approaches his craft with more care, patience, and skill than the amateur. Similar to the METAPHOR type, these descriptions lack specific subtypes as there were not enough distinguishable themes among the descriptions to justify clustering at a subtype level.

Despite the limitations of only having the bowl, the whisk, and the hand visible in the GIFs, respondents inferred the level of skill and effort that went into the motion from the GIFs. Respondents also gauged implicit attitudes from the motion, using words such as "deliberate", "purposeful", and "lazily", as if the motion was a reflection of the person performing the motion. This aligns with how people are able to gauge expertise and mastery based on a person's actions, e.g. being able to differentiate between lines made by a professional and a novice drawer.

6 CROWDSOURCING REFLECTIONS

We conducted our study on a crowdsourcing platform like Amazon Mechanical Turk (MTurk) to reach a large and varied population in a short amount of time. MTurk provides options to gauge worker quality based on metrics such as the number of HITs, or microtasks, completed by the crowdworker and their HIT approval rate. In this section, we reflect on how tweaking MTurk parameters affect the rate and quality of received responses, the cost-time balance for open-ended description tasks, and the scalability of the study based on the number of respondents and descriptions required to achieve the typology.

- **Approval Sweet Spot** Initially, we set a HIT approval rate to be greater than 70%. Although the typical HIT approval rate used for academic research is greater than 98%, we lowered the metric to explore if the lower barrier would yield a larger variety and invite more creativity in responses [29]. However, despite receiving responses in less than 24 hours, most responses only had completed the informed consent. We increased the approval rate to 90% and added a Master qualification (an MTurk qualification award). We received an initial influx of 10 respondents in the first 24 hours, yet

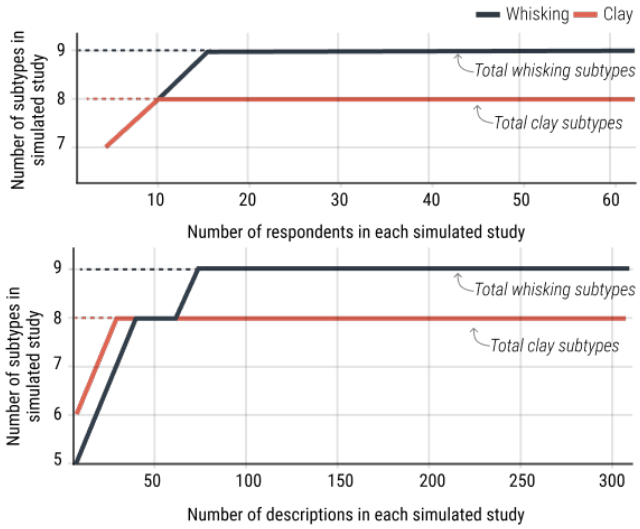

Figure 5: *Subtype Saturation* represents the average number of TDT subtypes present in each simulated study. The graphs show subtype saturation for 100 simulated crowdstudies.

this effect was temporary for the rest of the batch likely due to the relatively few crowdworkers that have the Master qualification. For the final iteration, we set the HIT approval rate to 98% and collected responses from 60 respondents in batches of different sizes, each taking under 30 hours. We pooled the responses from the second and third iterations for our results.

- **Cost-Time Balance** The average completion time per survey (which consisted of 5 whisking questions and 5 clay questions) was approximately 10 minutes. For the first two iterations, crowdworkers were compensated $0.50 (USD) for completing the survey. To increase the incentive to complete the survey and make the compensation more commensurate with the time required, we increased the compensation to $1.75 (USD) for completing the survey. These rates align with compensation rates used by other researchers to conduct ethical research on the crowdworking platform[29].
- **Subtype Saturation** While we did encounter diversity in the responses from crowdworkers, we did reach a point where we reached saturation. For applying our crowdsourcing method to other tasks, it is important to understand just how many descriptions or users are needed to reach comparable results. Although not exhaustive, we consider the TDT as a starting point for understanding the types of descriptions needed to capture the tacit and ambiguous information present in a task.
  Using a Monte-Carlo simulation approach, we created simulated studies by randomly sampling crowdworkers or descriptions from either the clay or whisk study corpus. This enabled us to create simulated studies with different sample sizes. We repeated each simulation 100 times; in this way, we could understand the probability of different types of descriptions occurring given specific study parameters and served as a time and cost-efficient alternative to running multiple HITs in MTurk.
  First, we report the *subtype saturation*, or average number of TDT subtypes represented in each simulated study, for simulated studies with 5 to 60 workers (Figure 5). Our findings indicate that subtype saturation was reached when simulated studies had 20 respondents. This sample size is a rather small and achievable sample size that is often seen in in-lab studies. This sample size could greatly be reduced if leaving out or targeting rarer subtypes (e.g., SOUND) in studies.
  Second, we report subtype saturation for simulated studies with random samples of descriptions from the cleaned description corpus of either task. We incrementally increased the sample size from 10 descriptions to the total number of descriptions for the specific task (294 for clay, 315 for whisk). For clay descriptions (uncommon experience), subtype saturation was reached before 50 descriptions whereas for the whisking descriptions (common task), subtype saturation was reached at around 70 descriptions.

7 DESIGN CONCEPTS

We propose three design concepts to illustrate how TDT can be leveraged in designing systems where tacit knowledge is prevalent. The concepts were motivated by prior research on understanding how tutorial systems and instructional materials could be enhanced to better communicate tacit and experiential knowledge [9]. We position these concepts in exemplar scenarios within creative practices where there is an expert and an audience engaged in a creative task and explore how these concepts translate across different mediums through which experiences are communicated. We describe the design concepts in the subsections below; we annoate examples of tacit descriptions used in the concepts with their typological relationships using TYPE > SUBTYPE notation.

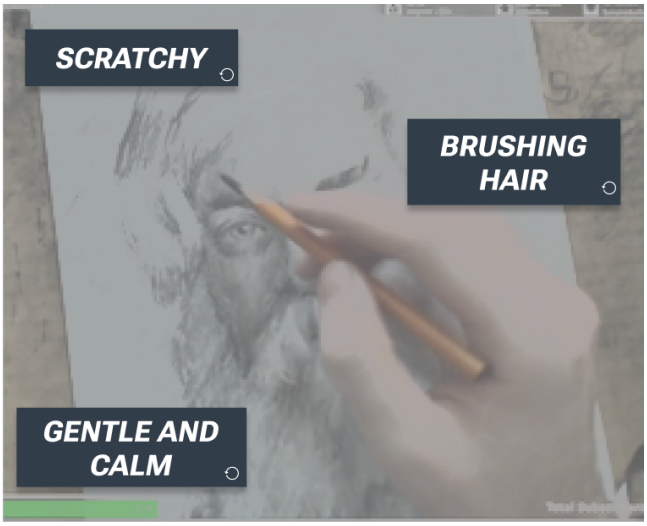

Figure 6: Tacit Layer. Tacit Layer communicates experiential and tacit knowledge that the streamer does not verbalize through dynamic annotation.

7.1 Tacit Layer

A tacit layer describes an interaction that prioritizes a breadth-first interpretation of information. Within tutorials, tacit layers inject information into the user's view that is relevant to the task at hand; unlike a single label or annotation, a tacit layer aims to provide a complete and comprehensive understanding of the task at hand. When applied to the Tacit Description Typology, a tacit layer can satisfy breadth-first interpretation by spanning all types of descriptions.

Consider a creative live stream of a drawing process (Figure 6). The live stream medium inherently constrains a streamer to narrate over their actions (INSTRUCTION) while the viewer is focused on the visuals being produced by the streamer (EXPERIENCE > VISUAL). The streamer or audience could supplement their view by superimposing other types of descriptions. An EXPERIENCE > SOUND annotation, such as "scratchy" can communicate to the audience unspoken details of the pressure, stroke length, and rhythm of the interaction (INSTRUCTION). A CARE annotation "gentle and calm" helps the audience consider the level of attention they should be exuding when using this drawing technique. To match the liveness of the interaction, we view these annotations updating, like a ticker tape, sampling the diverse descriptions of the corpus. While the "scratchy" label may not resonate with every member, it allows the right annotation (e.g., "counting pinto beans" or "brushing hair" - METAPHOR) to resonate with different audience members and communicate even more tacit information (i.e., only apply a pull stroke).

7.2 Tacit Space

A tacit space describes an interaction that prioritizes a depth-first interpretation of a single or isolated aspect of a task. In the context of creative tasks, a tacit space serves to create and demarcate the boundary between positive and negative actions. The space formed by these boundaries supports a user in understanding the multiple approaches to accomplishing the same task and choosing their own trajectory.

Consider the following steps for staining cells for flow cytometry.

Cell Staining Protocol

- Add 50 µL of whole blood to a 12 x 75 mm polystyrene tube.
- Add 20 µL fluorochrome-conjugated antibody to the sample solution.
- Vortex softly.
- …

A vortexer is a machine that mixes samples using a lateral, circular motion and, like a blender, has multiple settings. The ambiguity in the protocol above was likely written to provide flexibility to use the protocol with different vortexers. A tacit space can be formed to understand the Mixing Space of the vortexer by annotating the machine with lab-sourced tacit descriptions (Figure 7):

- Lvl. 1-3. Shakes, Barely Moves, Trembles, Mild, Quiet, Gentle, Touch It Barely, Stays in Lower Part of Tube
- Lvl. 4-6. Good Shake, Decent Mixing, Easy to Handle, Audible
- Lvl. 7-9. Very Shaky, Vigorous. Afraid Your Sample is Gonna Come Out, Bubbling Sample, Full Power, Only Hold for 5 ms, Hard to Hold Down, Loud Buzzing, Dancing on the Table

Although the descriptions easily help isolate the right setting for the "vortex softly" goal, the tacit space is also able to synchronously communicate the incorrect settings, or *negative spaces* of the vortexer, providing the biologist with a better understanding of how to make use of the vortexer's mixing capabilities.

In this way, tacit spaces can support users in constructing their own knowledge as a byproduct of working with materials and tools. For example, the tacit space of the vortexer may allow the biologist to surmise the necessary precautions for using a Level 9 mixing speed (i.e., using a larger tube and a smaller sample volume, locking the vortexer on the table) and design their own bioprotocols where that setting may be useful (i.e., homogenization, cell lysis, nanoparticle dispersion).

Figure 7: Tacit Space. A post-it captures the Mixing Space of the vortexer by annotating the machine with lab-sourced tacit descriptions.

7.3 Tacit Localizer

A tacit localizer serves as a contrast to learning approaches like cognitive apprenticeship that pair learners solely on expertise. Instead, a tacit localizer is a pairing interaction, applicable to either of the above concepts, that aims to help a user "find" themselves within the task or the practice by matching them to description contributors who see (perceive similar tacit qualities), think (hold alike mental models), or talk (generate similar semantics) like them. While there is value in pairing individuals by similarity, the opposite has also been shown to be beneficial for creative ideation [18].

In the context of a creative task, a tacit localizer can be used to filter the corpus of tacit descriptions through collaborative filtering (e.g., tasking the user with a tacit description task and matching on similar responses) or content-based filters (e.g., matching on user-contributor profile).

In the case of a line-dancing class (Figure 8), learners focus on different qualities when observing an expert, such as their expressions,

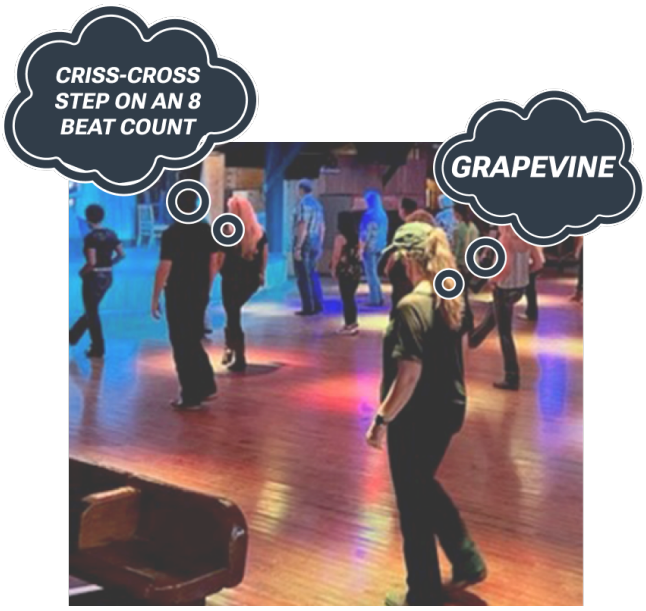

Figure 8: Tacit Localizer. Tacit Localizer "finds" a learner among their peers based on how they perceive a dance move.

movements, or timing. By having learners observe and describe a dance movement, instructors can generate more principled pairings. For example, while one learner may describe a motion as a "grapevine", another may describe it as a "criss-cross step on an 8-beat count". Pairing on tacit description types could serve as a way for instructors to scaffold and assess tacit skills (e.g., following beats-SOUND, understanding attention in line formations-CARE).

8 WORKSHOP

To ground our thinking of tacit descriptions, we conducted a workshop study. Inspired by card sorting techniques in design research [1, 10], we translated our design concepts into a set of card sorting activities for the workshop. Our goal was to gauge how participants consider tacit descriptions and how they envision our concepts in tutorial interactions in their respective tacit activities. Our goal was not to evaluate any functional prototypes or present the concepts in the specific scenarios they were grounded in.

8.1 Participants and Recruitment

We recruited six participants from our local network (Age 25.7±1.78, 4 Males, 2 Females). We publicized our study through word of mouth and sent out a pre-screening questionnaire to interested participants to get their contact information. Participants were invited to our lab for the workshop which took approximately 30 minutes. All participants were compensated with a 10 (USD) Amazon gift card. The study was approved by our institutional IRB.

8.2 Method

The workshop used a set of cards where each card consisted of a description sourced from the typology; we specifically used descriptions that came from the whisking task as these descriptions contributed to more subtypes in the typology than the clay descriptions. The suite of the card depicted the type of the typology and the color was determined by the whisking subtasks. In total, we had 38 cards containing descriptions that represented all the whisking subtypes in the typology as well as the five different GIFs (38 cards = 10 cards for task 1 sourced from subtask depicted in Figure 1 B5 + 24 cards for task 2 sourced from all subtasks + 4 cards for task 3 sourced from subtask depicted in Figure 1 B3). The cards were designed using FIGMA, a web-based design tool, and then printed on card stock paper.

Tasks. The workshop consisted of three tasks where participants were presented with cards and asked to perform the whisking motion represented by a set of cards. The whisking activity was only used as a means of introducing participants to the concepts before probing deeper for their insights. We did not judge participants based on how well they could replicate the whisking motion from the written descriptions. Although all three tasks involved whisking, the task activities and goals were different for each task, mitigating any learning effect that could occur from repeatedly whisking. We were more interested in observing participants' interactions with the cards and their responses to task specific interview questions on how they navigate tacit knowledge in their respective practices. For each task, participants were given a bowl with some heavy whipping cream, a hand-held balloon whisk, and a stack of cards. Example of a design card and the workshop setup is shown in Figure 9.

- Tacit Layer. We chose Figure 1 B5 (Supplemental material - whisk-5.gif) as the whisking motion to be replicated. The description corpus for this motion had at least one description from each typological type. Participants were given the cards face down and asked to progressively reveal the descriptions to incrementally develop their mental models. Our rationale was to observe how many descriptions participants went through to create a coherent perception of a motion. Participants were instructed to flip at least one card to reveal the description and whisk accordingly for 10 seconds. Participants had the choice to decide however many cards they wanted to flip to fully understand the motion but could not flip a card once they began whisking. After 10 seconds, we showed participants the original GIF of the whisking motion and asked how well they performed. We then conducted a brief interview where we asked participants about how they sought out information when learning a new task, how they navigated gaps in knowledge from tutorials, and how they wished information was presented to them instead.
- Tacit Space. We chose Figure 1 B1 (Supplemental material - whisk-1.gif) as the whisking motion to be replicated. For each whisking GIF, we extracted all descriptions that did not belong to the INSTRUCTION type and created five sets of design cards. We laid out the sets of cards in clusters in front of the participants. Participants were given some time to absorb the information. We explained how the different word clouds represented different whisking motions and each motion was different from the four other motions. Participants were then asked to whisk for 10 seconds after which we showed them the original GIF and conducted a

U1	A full-time software engineer. He is also a part-time pixel artist for an action-platformer Indie game.
U2	A graduate student in Computer Science. As a hobby, U2 enjoys cooking both for himself as well as friends and family.
U3	A graduate student in Computer Science. He has 8 and 4 years of experience playing badminton and pool respectively and currently mentors students in both sports.
U4	An undergraduate student in Computer Science. She likes to engage in adrenaline-inducing activities such as adventurous sports.
U5	An undergraduate student in Psychology. As a hobby, she likes to dance and create mixed-media art.
U6	A graduate student in Computer Engineering. He likes to play video games in his free time.

Table 1: Workshop User Profiles

brief interview. We asked participants about their affinity to learn from negative examples or from being told what not to do, the types of tasks where they would prefer more creative freedom, and how the boundaries of these tasks could be communicated to them.

- Tacit Localizer. Prior to the workshop, we wrote a Python script that counted the total frequency of each typology type from all descriptions provided by a respondent from the crowdsourcing study. The respondent was assigned a typology type that corresponded to the most recurring type in their descriptions. During the workshop, we first showed participants Figure 1 B2 (Supplementary Material whisk-2.gif) and asked them to write a description of the whisking motion in less than 5 words. One paper author immediately coded the description using the same coding scheme for the typology. We then showed the participant a description for Figure 1 B3 (Supplementary Material whisk-3.gif) that was provided by a respondent whose assigned type matched the participant's description type. For example, if the participant's description was coded with the EXPERIENCE type, we showed them a description from a respondent whose most recurring type was EXPERIENCE. We told participants that we were providing a description *from a person who thought like them*. Participants were then asked to whisk for 10 seconds based on this description. At the end of the task, we showed participants the GIF (Supplementary Material whisk-3.gif) and conducted a brief interview where we asked participants how they determined their and others' skill levels for different tasks within their creative practices and what factors influenced whom they liked to pair with for group tasks.

8.2.1 Participant Demographics. Since our local network consisted mostly of computer scientists and researchers, almost all participants were from Computer Science and Engineering backgrounds with the exception of one participant who was from a Psychology background. We were more interested in participants' side practices and hobbies which often included tacit activities. During our interviews, participants often applied our design concepts to their hobbies or side practices and in some instances, compared how the concepts might look different in their professions. To provide context to participant quotes and insights into the themes, we describe participants' professions and the specific tacit activities we discussed during their interviews in Table 1.

8.3 Themes

We noted general trends in participants' interactions with the cards and their interview responses during the interview. Post-interview, we analyzed video recordings of card activities and interviews to extract salient, insightful, and unexpected moments from which we generated memos. We conducted a thematic analysis on the memos leveraging Clarke and Braun's method [3, 7]. Both paper authors iteratively developed and applied inductive codes to the memos. The authors then discussed and refined the codes until full agreement was reached, and collaboratively clustered the codes into a set of descriptive themes. We present these themes below.

Seeking Meaning in Instructional Materials. In the Tacit Layer task, participants interacted differently with the TDT cards when trying to understand the motion to be replicated. U1, for example, immediately started categorizing the cards by placing them in clusters around the work table; these clusters mostly resembled the TDT subtypes, however, U1 had one cluster of descriptions that they felt was not specific enough. Some participants examined each card sequentially before putting them at the end of the deck, while others laid out the cards in a manner where they could glance at all of them as they were whisking. All participants went through the full set of cards, indicating a desire to know as much information about the motion before performing. Participants expressed different levels of satisfaction with how well they were able to replicate the motion once we showed them the GIF. When inquired about how they thought their motions were different, participants generally referred to a mismatch in their understanding of the speed and the direction of the motion. For example, U2 said they got confused about which direction was vertical in the circular bowl.

We asked participants if there was one card in the deck that truly captured the motions in the video. Popular responses were "splashing from left to right" (EXPERIENCE > SOUND, INSTRUCTION > DIRECTION) and "sloppy zig-zag" (CARE, INSTRUCTION > DIRECTION). On being asked to reflect why they chose "splashing from left to right", U5 said:

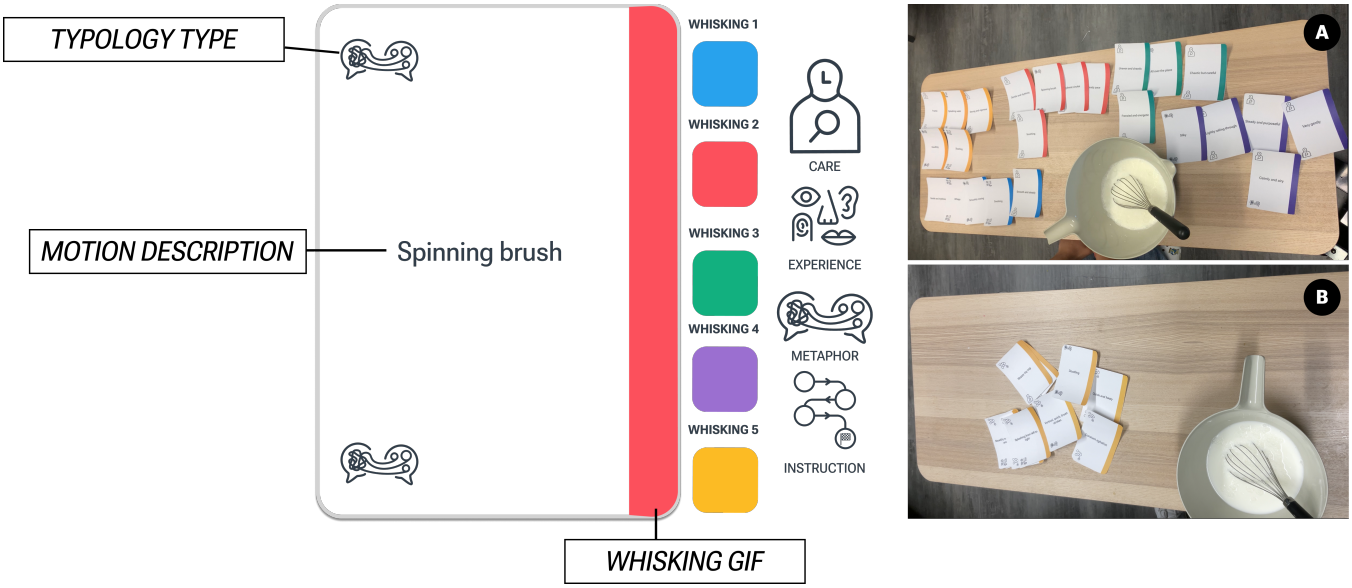

Figure 9: Anatomy of a Design Card. Each card shows a whisking motion description sourced from the typology. The suite of the card depicts the category of the typology and the color is determined by the whisking subtasks. A) Snapshot from a Tacit Space card activity. Similar colored cards were clustered together and all cards were laid out across the table to simulate the tacit space. B) Snapshot from a Tacit Layer card activity. The participant laid out the cards in a manner where they could glance at all of them as they were whisking.

> **U5** I think splashing captures how the whisk is held as well. If you are more sure about how you are holding it, you are actually whisking it. Versus if you are light-handed, you are splashing.

U5's rationale demonstrated how participants sought deeper meanings to words and related words to their experiences, such as trying to deduce the amount of care ("light-handed") that would be necessary to produce a sound ("splashing").

We further probed on what types of information participants sought out in instructional materials when learning a new task or a variation of a task and what information they felt was missing from these instructional materials. There was unanimous agreement that there was no singular tutorial that captured all information that participants needed. There was no expectation for such a tutorial either; as U5 said, it is "inevitable" that a lot of information is left out.

Participants expressed that instructional materials tended to leave out experiential information as a result of being too focused on technicalities. There was a desire to see their mental models being reflected in the tutorials, as participants wanted to hear from others "in their shoes" who had also interacted with the tutorials.

For example, when trying to learn new software, U2 had difficulty of finding a tutorial that aligned with his skill level.

> **U2** The examples the videos [tutorials] were showing me were not relevant. They were showing something very in-depth. I am just a learner so it [depth] was not helpful.

Similarly, U1 was frustrated when encountering documentation of programming libraries that are rife with jargon but leave out information about the creators' purpose or process.

> **U1** They don't really tell you much. [I want to know] What's the purpose these people made the function for, not what does this code do?

Additionally, U3 reported going through several hour-long tutorials when learning a new technical skill.

These insights indicate that while functional and technical instructions (INSTRUCTION) that tell participants what to do are useful in replicating a task, tacit instructions (EXPERIENCE, CARE, METAPHOR) are more meaningful for providing a holistic picture of the task and aid a learner in understanding how the task morphs for their specific contexts.

Creative Liberties in Tacit Spaces. During the Tacit Space task, some participants demonstrated a tendency to fixate only on the cards representing the motion to be replicated, sometimes only giving the other cards a cursory glance as they were being laid out. On probing further on why that was the case, participants alluded to wanting to focus on the task at hand.

> **U1** I did look at them but did not focus on them. Also, it's a lot easier for a person to think more on the lines of what to do versus what not to do.

> **U4** I think the spotlight was on that one [card clusters for target motion]. But I feel like if I had thought these [descriptions in other clusters] were things I shouldn't do, I would have whisked slower.

Following this line of thought, we explored participants' attitudes towards negative examples and their reactions to the creative freedom that comes with only being told what not to do. We asked them to think about how they would have whisked had they not

been given the descriptions of the target motion and brainstormed their experiences with negative examples in their daily tacit activities.

Participant responses were mixed. Some participants handled creative liberties with caution, unsure whether taking a leap of faith in a specific approach was worth failing to perform the task, underperforming, or causing harm to oneself or others. They also considered whether the time it takes to figure out what to do was worth the end goal, especially if they were not doing it right. U3 elaborated on this thought process when they were trying to learn a new move in pool on their own.

> **U1** If I try to do it on my own, even with experience, there are things you don't know. It can take me something like half an hour to figure it out and someone can show it to me in thirty seconds. On the other hand, I might never learn it actually even if I have infinite time if I am doing something wrong and I don't know that I am doing it wrong.

In contrast, U2 indicated feeling significantly limited by step-by-step instructions when cooking and preferred to work instead with a general idea of the cooking process.

> **U2** It's very difficult when you have all the individual ingredients listed in certain amounts. It's very complicated.

U2 mentioned how this mindset changed slightly when the stakes were higher, i.e. when they were cooking for friends or family, they were more careful with ingredient quantities.

U1 presented a unique perspective on negative examples, perceiving negative examples as a challenge to explore whether something good could come out of them.

> **U1** [In pixel art,] there are certain limitations on where you should put certain colors. It's saying that you shouldn't do this because it might look bad, however, there is always an instance that you might do it anyway and it still comes out good.

Participants' reactions to the simulated tacit space of the TDT cards, their reflections on creative liberties, and their attitudes towards negative examples suggest uncertainty in how to navigate this space. This may be motivated by habit or learning preferences which make the novelty of only being told what not to do seem intimidating. Additionally, having a lot of creative liberty only seem desirable in scenarios where the process outweighs the outcome for the individual participant.

Finding yourself in a creative space. In the Tacit Localizer task, participants interpreted the concept of receiving a description from "a person who thinks like them" differently. For example, U2 interpreted it as an opportunity to include his own whisking experiences; whereas the description asked to whisk "around and across", U2 only whisked around the bowl and not across.

U4 received the same description and whisked in both directions. However, upon seeing the video from which the description was sourced, U4 commented on how the description did not describe the bowl being tilted, indicating that she could have caught that detail in the description.

U5 attributed the difference in their whisking motion to that of the videos due to the material appearing different (The cream in the video was a lot thicker than the one we used), alluding that if the setup was exactly the same as the video, they might have been able to replicate better.

We probed further on how participants tried to find "like-minded" people when trying to perform a collaborative task. We observed a general trend of responses hinging around expertise, with several participants saying they liked working with others with similar expertise and thought processes who can relate to their struggles. Some participants expressed curiosity in wanting to be paired with someone with different expertise to learn alternative perspectives, but this opinion had relatively less enthusiasm. Talking about multiplayer games, U6 summed up the sentiments by saying:

> **U6** The benefit of working with someone with a similar mindset is that you can work as a team. But with a different mindset, you can learn more.

However, U2 was adamant about not wanting to work with someone with a large expertise gap, stating:

> **U2** I feel like embarrassed that if I ask a question, that person may feel like oh he is stupid.

We also wanted to understand how a mentor might group learners. U3, who mentors novice badminton and pool players, described how he organizes teams for practice as follows:

> **U3** The pairing works mostly better if there is a skill gap. I would pair people who are beginners with people who know how to play.

When further probed about if he would consider pairing mentees based on their thinking, he said:

> **U3** It could work but it would be difficult to get mindset because I don't think it will always be 100% correct.

Participants' card interactions and responses demonstrated the ambiguity in situating oneself within a corpus for a specific task. Participants' interpretation of a like-minded individual when they were provided with the information was more nuanced than their interpretation when they sought after a like-minded individual. When they were provided with the information, participants interpreted like-mindedness as similarities in experiences and perceptions, whereas when seeking a like-minded individual, they only considered expertise.

9 DISCUSSION

We leverage our findings from the workshop to discuss future research trajectories for tacit descriptions. We provide design implications for how tacit descriptions can be encoded in different mediums, describe its potential in addressing the challenges of tacit knowledge communication through semantic vocabularies, and discuss the possibility of integrating tacit descriptors into psychometric scales.

9.1 Tacit Descriptions in Alternate Mediums

Our workshop leveraged design cards with written descriptions as the medium for encoding tacit descriptions, enabling participants to interact with the text in a physical form. Exposing participants to too many written descriptions in an attempt to communicate the tacit space proved less effective as participants only focused on the target descriptions, demonstrating a need to encode tacit descriptions in a way that is not cognitively overwhelming. Encoding

tacit information in alternative mediums would require a careful negotiation between the medium's affordances and the amount of information to be encoded.

The amount of tacit encodings to provide is primarily motivated by the audience's need. Findings from the Tacit Layer task suggest that tacit encodings should complement tutorial instructions, not repeat them. In particular, emphasis should be put on experiential and sensory information that is typically missing from tutorials.

Information density can also vary if the natural affordances of the medium contribute to the sensemaking process. For example, the omnidirectional nature of 360-degree videos presents an exciting opportunity to simulate a tacit space while also reducing the amount of sensemaking required by providing more visual cues. Recent works in immersive audio descriptions show promise in the audio being a lucrative medium for encoding the descriptions in these videos [5]. Alternatively, there is also room to explore how tacit descriptions can be rendered in mediums that have traditionally supported tacitness while being cognitively digestible. For example, the pace, direction, and handedness of a whisking motion could also be communicated through expressive line-drawing. Additionally, we see an exciting new frontier in encoding tacit descriptions through tangible and embodied interactions, which can leverage the body's natural senses in capturing more meaningful information.

9.2 Generating Semantic Vocabularies

We extracted tacit descriptions through text format which is a powerful medium for making explicit the semantics of tacit activities. People's choice of words carries implicit meaning; tacit descriptions combine this inherent meaning with experiential, cultural, and context cues, making them semantically richer. For example, a noteworthy description of whisking motion was "bitch slapping motion". Although at first glance the description might seem inappropriate, we argue that it is perhaps one of the most semantically rich and interesting descriptions in our typology as it communicates pace, motion, direction, and force (INSTRUCTION) through a metaphor (METAPHOR) but also encapsulates socio-cultural nuances. In the same way the meaning of words morph through generations and our vocabularies reflect the times we have lived through (e.g. "tea" is slang used by millennials to mean gossip), tacit descriptions are also reflections of the people who have experienced the activity. We briefly leverage this idea in the Tacit Localizer concept and see an even greater opportunity in applying the idea to the larger problem of tacit communication not having a universal vocabulary. Prior works have shown how semantic labels could be leveraged as a sense-making tool to help users explain why they like an image [20]. Similarly, tacit descriptions can aid in the development of semantic vocabularies to describe tacit information. One potential way to aid in developing this vocabulary could be through introducing tacit descriptions as captions on tutorial videos; captions have proven to be effective in helping people learn languages [13].

9.3 Use of Experiential Units in Psychometrics

While processing descriptions for TDT, we were curious to see if participants used formal or scientific units of measurement, such as an SI unit, to describe the qualities of the motion. We found only one participant who described the pace of the motion in revolutions per second. Instead, participants leveraged what we coin *experiential units*, which are informed by collective experiences. For example, the word "fast" does not convey a singular velocity, yet people have a general idea of what constitutes a fast pace.

Experiential units are inherently ambiguous in nature; in the Tacit Layer task, participants' motions often were visibly different from each other as well as that depicted in the target video despite following the same instructions. However, since these types of units are a natural part of our everyday vocabulary, which extends to popular learning tools such as tutorials, we see great value in the existence of a psychometric scale that captures the ranges of each experiential unit. In psychometrics, semantically differential scales already demonstrate their benefit over traditional Likert scales by capturing implicit and complex attitudes. An experiential scale derived from tacit descriptions can potentially capture even more nuances and is directly applicable to a wide range of practices that involve tacit communication (e.g., torque communicated through anchors like "monkey tight" and "gorilla tight".) We envision such a scale to be particularly useful in interpreting tutorial artifacts which are rife with experiential units but often require the learner to closely follow the expert's motion.

9.4 Limitations

We chose to use whisking and clay GIFs in our crowdsourcing tasks because the whisking motion is familiar to many, while clay consistencies are unfamiliar to those outside of the ceramics practice. However, respondents may have encountered clay or clay-like materials before, making it easier to communicate their perceptions. Tacit activities that require niche domain-specific knowledge to fully experience the activity can be challenging for crowdworkers to describe as they may lack the expertise to provide detailed descriptions. This limits the range of possible tasks from which tacit information can be extracted using the crowdsourcing method. However, our subtype saturation investigation showed that even a small group of crowdworkers can provide rich and detailed tacit descriptions. Since crowdworkers have diverse backgrounds, it could still be worthwhile to use a crowdworking platform for domain specific tasks by filtering for participants who have experience with the tasks. Additionally, our tasks used very short GIFs and limited the size of the expected descriptions to less than five words. However, not all tasks can be isolated into GIFs and would require showing a longer process to engender meaningful descriptions. In future work, we aim to explore how such tasks could be designed to make them fit the standards of a crowdsourcing task while also making it possible to extract rich tacit descriptions.

10 CONCLUSION

In this work, we presented a crowd-based method for extracting tacit descriptions from short tutorial GIFs of familiar motions and unfamiliar materials. We synthesized the crowdsourced descriptions into a Tacit Description Typology that categorized these descriptions into four types according to how they communicated explicit and tacit information. We leveraged the typology in the design of three concepts that demonstrated how tacit descriptions can enhance tutorial artifacts. We then conducted a workshop study

where we translated our design concepts into card-sorting activities that gauged how participants consider tacit descriptions in their respective practices. Our findings have implications for the design of systems where tacit and experiential knowledge are prevalent.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for their insightful feedback, the peer support from The Hybrid Atelier at UTA, and our research participants, without whom this project would not have been possible. This research was funded by the National Science Foundation Grant #2105054.

REFERENCES

[1] Tessa Aarts, Linas K Gabrielaitis, Lianne C De Jong, Renee Noortman, Emma M Van Zoelen, Sophia Kotea, Silvia Cazacu, Lesley L Lock, and Panos Markopoulos. 2020. Design card sets: Systematic literature survey and card sorting study. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 419–428.

[2] Jordan Beck and Hamid R Ekbia. 2018. The theory-practice gap as generative metaphor. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–11.

[3] Virginia Braun and Victoria Clarke. 2021. Thematic analysis. Analysing qualitative data in psychology.

[4] Yuanzhi Cao, Anna Fuste, and Valentin Heun. 2022. MobileTutAR: a Lightweight Augmented Reality Tutorial System using Spatially Situated Human Segmentation Videos. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3491101.3519639

[5] Ruei-Che Chang, Chao-Hsien Ting, Chia-Sheng Hung, Wan-Chen Lee, Liang-Jin Chen, Yu-Tzu Chao, Bing-Yu Chen, and Anhong Guo. 2022. OmniScribe: Authoring Immersive Audio Descriptions for 360 Videos. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–14.

[6] Pei-Yu Chi, Joyce Liu, Jason Linder, Mira Dontcheva, Wilmot Li, and Bjoern Hartmann. 2013. Democut: generating concise instructional videos for physical demonstrations. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*. 141–150.

[7] Victoria Clarke, Virginia Braun, and Nikki Hayfield. 2015. Thematic analysis. *Qualitative psychology: A practical guide to research methods* 222, 2015 (2015), 248.

[8] John S Edwards. 2003. Knowledge engineering: A forgotten element in knowledge management. In *Keynote Presentation in the Operational Research Society Annual Conference. Retrieved October*, Vol. 12. 2004.

[9] Shreyosi Endow and Cesar Torres. 2021. "I'm Better Off on my Own": Understanding How a Tutorial's Medium Affects Physical Skill Development. In *Designing Interactive Systems Conference 2021*. 1313–1323.

[10] Angi Faiks and Nancy Hyland. 2000. Gaining user insight: a case study illustrating the card sort technique. *College & research libraries* 61, 4 (2000), 349–357.

[11] Adam Fourney and Michael Terry. 2014. Mining online software tutorials: challenges and open problems. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. Association for Computing Machinery, New York, NY, USA, 653–664. https://doi.org/10.1145/2559206.2578862

[12] Florian Gadsby. 2021. How I reclaim, cut wedge and spiral wedge my clay — narrated video. https://www.youtube.com/watch?v=mco55yi_U6s.

[13] Susan Gass, Paula Winke, Daniel R Isbell, and Jieun Ahn. 2019. How captions help people learn languages: A working-memory, eye-tracking study. (2019).

[14] Camilla Groth, Maarit Mäkelä, and Pirita Seitamaa-Hakkarainen. 2015. Tactile augmentation: A multimethod for capturing experiential knowledge. *Craft Research* 6, 1 (2015), 57–81.

[15] Kotaro Hara, Abigail Adams, Kristy Milland, Saiph Savage, Benjamin V Hanrahan, Jeffrey P Bigham, and Chris Callison-Burch. 2019. Worker demographics and earnings on amazon mechanical turk: An exploratory analysis. In *Extended abstracts of the 2019 chi conference on human factors in computing systems*. 1–6.

[16] Panagiotis G Ipeirotis. 2010. Demographics of mechanical turk. (2010).

[17] Lee Jones and Audrey Girouard. 2022. Learning with Stitch Samplers: Exploring Stitch Samplers as Contextual Instructions for E-Textile Tutorials. In *Designing Interactive Systems Conference* (Virtual Event, Australia) *(DIS '22)*. Association for Computing Machinery, New York, NY, USA, 949–965. https://doi.org/10.1145/3532106.3533488

[18] Andruid Kerne, Andrew M Webb, Steven M Smith, Rhema Linder, Nic Lupfer, Yin Qu, Jon Moeller, and Sashikanth Damaraju. 2014. Using metrics of curation to evaluate information-based ideation. *ACM Transactions on Computer-Human Interaction (ToCHI)* 21, 3 (2014), 1–48.

[19] America's Test Kitchen. 2014. Science: The Best Way to Use a Whisk (Have You Been Whisking All Wrong?). https://www.youtube.com/watch?v=zglSRFlFHs&t=3s.

[20] Janin Koch, Nicolas Taffin, Andrés Lucero, and Wendy E Mackay. 2020. SemanticCollage: Enriching digital mood board design with semantic labels. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 407–418.

[21] Benjamin Lafreniere, Tovi Grossman, and George Fitzmaurice. 2013. Community enhanced tutorials: improving tutorials with multiple demonstrations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 1779–1788. https://doi.org/10.1145/2470654.2466235

[22] Gierad Laput, Eytan Adar, Mira Dontcheva, and Wilmot Li. 2012. Tutorial-based interfaces for cloud-enabled applications. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*. Association for Computing Machinery, New York, NY, USA, 113–122. https://doi.org/10.1145/2380116.2380132

[23] Panagiotis Lazanas. 2012. Subconscious (tacit) structures of expertise.

[24] Yuncheng Li, Yale Song, Liangliang Cao, Joel Tetreault, Larry Goldberg, Alejandro Jaimes, and Jiebo Luo. 2016. TGIF: A new dataset and benchmark on animated GIF description. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 4641–4650.

[25] Lucas M. Silva and Daniel A. Epstein. 2021. Investigating preferred food description practices in digital food journaling. In *Designing Interactive Systems Conference 2021*. 589–605.

[26] Nada Makes. 2017. How to Recycle and Dry Clay to the Right Consistency. https://www.youtube.com/watch?v=qDEa31gsAHA&t=1s.

[27] David Martin, Benjamin V Hanrahan, Jacki O'Neill, and Neha Gupta. 2014. Being a turker. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 224–235.

[28] Peter Mohr, David Mandl, Markus Tatzgern, Eduardo Veas, Dieter Schmalstieg, and Denis Kalkofen. 2017. Retargeting Video Tutorials Showing Tools With Surface Contact to Augmented Reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 6547–6558. https://doi.org/10.1145/3025453.3025688

[29] Jonas Oppenlaender, Kristy Milland, Aku Visuri, Panos Ipeirotis, and Simo Hosio. 2020. Creativity on paid crowdsourcing platforms. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[30] Michael Polanyi. 2009. The tacit dimension. In *Knowledge in organizations*. Routledge, 135–146.

[31] Wendy Ross and Mike Groves. 2022. Let's Just See What Happens: A Qualitative Case Study of Risk and Uncertainty in the Creative Process. *The Journal of Creative Behavior* (2022).

[32] Richard Sennett. 2008. *The craftsman*. Yale University Press.

[33] Thomas Smith, Simon Bowen, Bettina Nissen, Jonathan Hook, Arno Verhoeven, John Bowers, Peter Wright, and Patrick Olivier. 2015. Exploring Gesture Sonification to Support Reflective Craft Practice. *null* (2015). https://doi.org/10.1145/2702123.2702497

[34] Cesar Torres, Matthew Jörke, Emily Hill, and Eric Paulos. 2019. Hybrid Microgenetic Analysis: Using Activity Codebooks to Identify and Characterize Creative Process. In *Proceedings of the 2019 on Creativity and Cognition*. 2–14.

[35] Anh Truong, Peggy Chi, David Salesin, Irfan Essa, and Maneesh Agrawala. 2021. Automatic Generation of Two-Level Hierarchical Tutorials from Instructional Makeup Videos. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3411764.3445721

[36] Nicola Wood. 2008. The new journeyman: The role of an expert learner in eliciting and transmitting skilled knowledge. (2008).

[37] Nicola Wood, Chris Rust, and Grace Horne. 2009. A tacit understanding: The designer's role in capturing and passing on the skilled knowledge of master craftsmen. *International Journal of Design* 3, 3 (2009).

[38] Masahiro Yamaguchi, Shohei Mori, Peter Mohr, Markus Tatzgern, Ana Stanescu, Hideo Saito, and Denis Kalkofen. 2020. Video-Annotated Augmented Reality Assembly Tutorials. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. Association for Computing Machinery, New York, NY, USA, 1010–1022. https://doi.org/10.1145/3379337.3415819

[39] Mingrui Ray Zhang, Mingyuan Zhong, and Jacob O Wobbrock. 2022. Ga11y: An Automated GIF Annotation System for Visually Impaired Users. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–16.

[40] Mingyuan Zhong, Gang Li, Peggy Chi, and Yang Li. 2021. HelpViz: Automatic Generation of Contextual Visual Mobile Tutorials from Text-Based Instructions. In *The 34th Annual ACM Symposium on User Interface Software and Technology (UIST '21)*. Association for Computing Machinery, New York, NY, USA, 1144–1153. https://doi.org/10.1145/3472749.3474812